Rapid perceptual integration of facial expression and emotional body language

Hanneke K. M. Meeren*†, Corné C. R. J. van Heijnsbergen*, and Beatrice de Gelder*†‡§

*Cognitive and Affective Neuroscience Laboratory, Tilburg University, LE 5000, Tilburg, The Netherlands; †F. C. Donders Centre for Cognitive Neuroimaging, 6500 HB Nijmegen, The Netherlands; and ‡Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129

In our natural world, a face is usually encountered not as an isolated object but as an integrated part of a whole body. The face and the body both normally contribute in conveying the emotional state of the individual. Here we show that observers judging a facial expression are strongly influenced by emotional body language. Photographs of fearful and angry faces and bodies were used to create face–body compound images, with either matched or mismatched emotional expressions. When face and body convey conflicting emotional information, judgment of facial expression is hampered and becomes biased toward the emotion expressed by the body. Electrical brain activity was recorded from the scalp while subjects attended to the face and judged its emotional expression. An enhancement of the occipital P1 component as early as 115 ms after presentation onset points to the existence of a rapid neural mechanism sensitive to the degree of agreement between simultaneously presented facial and bodily emotional expressions, even when the latter are unattended.

emotion communication | event-related potentials | visual perception

The face and the body both normally contribute in conveying the emotional state of the individual. Darwin (1) was the first to describe in detail the specific facial and bodily expressions associated with emotions in animals and humans and regarded these expressions as part of emotion-specific adaptive actions. From this vantage point, face and body are part of an integrated whole. Indeed, in our natural world, faces are usually encountered not as isolated objects but as an integrated part of a whole body. Rapid detection of inconsistencies between them is beneficial when rapid adaptive action is required from the observer. To date, there has been no systematic investigation into how facial expressions and emotional body language interact in human observers, and the underlying neural mechanisms are unknown.

Here, we investigate the influence of emotional body language on the perception of facial expression. We collected behavioral data and simultaneously measured electrical event-related potentials (ERP) from the scalp to explore the time course of neuronal processing with a resolution in the order of milliseconds. We predicted that recognition of facial expressions is influenced by concurrently presented emotional body language, and that affective information from the face and the body start to interact rapidly. In view of the important adaptive function of perceiving emotional states and previous findings about rapid recognition of emotional signals, we deemed it unlikely that integrated perception of face–body images results from relatively late and slow semantic processes. We hypothesize that the integration of affective information from a facial expression and the accompanying emotional body language is a mandatory automatic process occurring early in the processing stream, which does not require selective attention, thorough visual analysis of individual features, or conscious deliberate evaluation.

Photographs of fearful and angry faces and bodies were used to create realistic-looking face–body compound images with

either matched or mismatched emotional expressions. Electroencephalogram (EEG) signals were recorded while subjects attended to the face and judged the emotional expression. A short stimulus-presentation time was used (200 ms), requiring observers to judge the faces on the basis of a "first impression" and to rely on global processing rather than on extensive analysis of separate facial features. The emotions fear and anger were selected because they are both emotions with a negative valence, and each is associated with evolutionarily relevant threat situations. Although subcortical circuitry may be capable of some qualitative differentiation between fear and anger (2), their differentiation in terms of cortical responses is as minimal as possible. Differential ERP responses have not been reported to occur before ≈170 ms (3). This will make observed differences between electrophysiological correlates of congruent and incongruent stimuli easier to interpret.

To test our automatic-processing hypothesis, we focused on the analysis of early electrophysiological components that can be readily identified in the visual ERP waveform, i.e., the P1 and the N170 component. The N170 is typically found in face perception studies as a prominent negative brain potential peaking between 140 and 230 ms after stimulus onset at lateral occipito-temporal sites. It shows a strong face sensitivity putatively pointing to specialized neuronal processing routines for faces as compared with objects. It is commonly considered as an index of the early stages of encoding of facial features and configurations (e.g., ref. 4). The waveform shows a robust face-selective "inversion" effect indicative of configural processing; i.e., it is enhanced and delayed to faces that are presented upside down but not to inverted objects (5, 6). In general, it has been found to be insensitive to aspects of the face other than the overall structure, like expression or familiarity. Of interest, a typical but slightly faster N170 component showing the typical inversion effect commonly obtained for faces was also found for the perception of human bodies (6).

Recent studies have challenged the N170 as the earliest marker of selective face processing and drawn attention to an earlier component. The P1 component is a positive deflection recognizable at occipital electrodes with an onset latency between 65 and 80 ms and peaking around 100–130 ms. There is evidence that this component is mainly generated in "early" extrastriate visual areas (e.g., refs. 7–10), and it is commonly thought that it reflects processing of the low-level features of a stimulus. Its amplitude is modulated by spatial attention, but selected attention to nonspatial features does not affect the component (e.g., ref. 11). However, a few recent studies suggest that when it comes to faces, at least some form of higher-order processing can already occur at this early stage. Using magnetoencephalography, a face-selective component has already been identified around 100–120 ms after stimulus onset that was

found to correlate with successful categorization of the stimuli as faces (7, 12). In addition, both the magnetic and electrical components have been found to be sensitive for face inversion (7, 13, 14), suggesting that some configurational processing already takes place at this stage. Although most ERP studies only reveal consistent differences between emotional and neutral faces after 250 ms and among emotions after 450 ms, there is some evidence that the P1 may be sensitive to affective information contained in the face (15). A few studies have reported a global effect of facial expression on the P1 (e.g., refs. 3, 16, and 17), although differences between distinctive emotions have not been found (3, 18).

Materials and Methods

Participants. Twelve healthy right-handed individuals (mean age 20 years, range 18–24 years; three males) with normal or corrected-to-normal vision volunteered after giving their consent to take part in the study. The study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

Stimulus Material. The body stimuli were taken from our own validated data set and have previously been used in behavioral, EEG (6), and functional MRI studies (e.g., refs. 19 and 20). Photographs consisted of gray-scale images of whole bodies (four males and four females) adopting a fearful posture and an angry posture (for details, see ref. 20). In both the fearful and angry body poses, there was a large variability in the position of the arms, which ranged from slightly to fully bended. The hands were usually extended, with hand palms pointing outward in the fear bodies (defensive pose) while being flexed and pointing inward in the angry body poses. Face stimuli were gray-scale photographs from the Ekman and Friesen database (21). Eight identities (four male, four females) were used, each with a fearful and an angry expression.

Face–body compound stimuli were created by using photo-editing software. The faces and bodies were cut out and carefully sized and combined in realistic proportions (face–body ratio of ≈1:7) to create eight new "identities" (four males, four females). Four categories of compound stimuli were made, each containing the eight different identities: two categories of "congruent" compounds with matching facial and bodily expressions (i.e., fear and anger) and two categories of "incongruent" compounds for which facial and bodily expression did not match (i.e., fearful faces combined with angry bodies and angry faces combined with fearful bodies). Examples of these stimulus categories can be found in Fig. 1.

The same faces and bodies were also presented in isolation, thus serving as control stimuli. The isolated stimuli preserved the same size and position on the screen as in the "original" compound stimuli. In the isolated body figures, only the outer contour of the head was visible; heads were "filled" with gray background color. The isolated face stimuli consisted of heads only; no outlines of the bodies were visible.

Experimental Procedures. The study comprised two experiments: the Congruence experiment during which the face–body compound stimuli were presented and a control experiment during which the same participants viewed pictures of the isolated faces and isolated bodies. Each experiment was divided into two blocks consisting of 128 trials. Within one block, all stimuli were presented four times in random order, hence summing up to a total of 64 trials per stimulus category for each experiment. Congruence and control blocks alternated. The temporal order of the blocks was counterbalanced across participants: half of the subjects started with a Congruence block, the other half with a control block. To familiarize the subjects with the procedure and

Fig. 1. Examples of the four different categories of Face–Body Compound stimuli used. Congruent and incongruent stimuli consisted of exactly the same material. The bodies of the two congruent stimulus conditions were swapped to create a mismatch between the emotion expressed by the face and the body.

task demands, the experiment was preceded by a short training session, which contained all stimulus categories.

The experiments were conducted in an electrically shielded sound-attenuating room. Subjects were comfortably seated in a chair with their eyes at 80-cm distance from the screen. The size of the framed face–body compound stimuli on the screen was 3.9×7.9 cm ($=1.4°$ horizontal $\times 2.8°$ vertical visual angle). Each trial started with a 1,000-ms fixation point (white on a black background) slightly above the center of the screen, i.e., at "breast height" of the depicted bodies in the images. Hence, the face and upper body (including the hands) subtended an equally large visual angle (equal eccentricity of critical facial and body features), which greatly minimized the tendency to make eye movements. Stimuli were presented for 200 ms and were followed by a black screen. The participant's task was to keep his or her eyes fixed on the fixation point and to decide as accurately and rapidly as possible whether the stimulus was expressing fear or anger by using two designated buttons (right-hand responses). Participants were explicitly instructed to judge the expression of the face while viewing the face–body compound stimuli. During the presentation of the isolated stimuli, they had to judge the expression conveyed by the presented stimulus, which could either be a face or a body. The next stimulus followed at 1,000 ms after the response.

EEG Recording. EEG was recorded from 49 locations by using active Ag-AgCl electrodes (BioSemi Active-Two, BioSemi, Amsterdam) mounted in an elastic cap, referenced to an additional active electrode (Common Mode Sense). EEG signals were band-pass-filtered (0.1–30 Hz, 24 dB/octave) and digitized at a sample rate of 256 Hz. In addition, horizontal and vertical electro-oculogram (EOG) were registered to monitor eye movements and eye blinks.

Data Analysis. Off-line, the raw EEG data were rereferenced to an averaged reference and were segmented into epochs starting 100 ms before to 1,000 ms after stimulus onset. The average amplitude of the 100-ms prestimulus epoch served as baseline. The data were electro-oculogram (EOG)-corrected by using the algorithm of Gratton *et al.* (22). Trials with an amplitude change exceeding 100 μV at any channel after EOG correction were rejected from analysis. Signals were averaged across trials time-

Table 1. Behavioral results, face–body compound stimuli

EmoFace	EmoBody	Congruence[A,R]	Accuracy, %	Reaction time, ms
Anger	Anger	Congruent	83.6 ± 2.6****a	778 ± 52***a
	Fear	Incongruent	70.6 ± 3.6****b	835 ± 57***b
Fear	Fear	Congruent	78.8 ± 3.3****a	771 ± 57***a
	Anger	Incongruent	63.9 ± 5.5****b	845 ± 49***b

Mean ± SEM values for accuracy and reaction time during the emotional expression judgment task and the results of the GLM: A and R in column heading denote main effects for accuracy and reaction time, respectively, for a given factor. Asterisks denote corresponding P values (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; and ****, $P < 0.0001$), with a and b indicating contrasting conditions. Group data of 12 participants.

locked to the onset of the pictures separately for each stimulus category to create the following category-specific ERPs: four different compound categories, FearFace–FearBody, FearFace–AngryBody, AngryFace–FearBody, and AngryFace–AngryBody; and four different categories for the control stimuli, i.e., the isolated faces and bodies: FearFace, AngryFace, FearBody, and AngryBody. The P1 component was identified at occipital sites (O1, Oz, and O2) as the first prominent (with amplitude higher than two times the maximal prestimulus deflection) positive deflection appearing between 80 and 140 ms after onset of the picture. The N170 component was identified at the occipito-temporal sites P7 and P8 as the maximal negative peak in the time window 140–230 ms. Peak amplitude and latency were scored for these components and subjected to analyses of variance for repeated measures using the General Linear Model. The following factors were used as within-subject factors for the face–body compound stimuli: Leads (two or three levels), EmoFace (i.e., facial expression: fear, anger), and Congruence (congruent, incongruent). Note that any given EmoFace–Congruence interaction equals an EmoBody main effect. Leads, StimulusType (face, body), and Emotion (fear, anger) were used for the control stimuli. When appropriate, post hoc tests were performed and their P values were Bonferroni-corrected. The parameters Accuracy (no. of correct responses/ no. of total responses × 100%) and Reaction Time of the behavioral responses were analyzed in a similar fashion, with the factor Leads omitted.

Results

Behavioral Results. Table 1 and Fig. 2 show the behavioral performance on the facial expression judgment task for the compound stimuli. A main effect of Congruence for both accuracy [$F(1, 11) = 36.04$, $P < 0.0001$] and reaction time [RT; $F(1, 11) = 27.95$, $P < 0.001$] was found. Participants made significantly better (accuracy of 81%) and faster (RT of 774 ms) decisions when faces were accompanied by a matching bodily expression than when a bodily expression did not match the facial expression (67% and 840 ms). This effect of congruence ob-
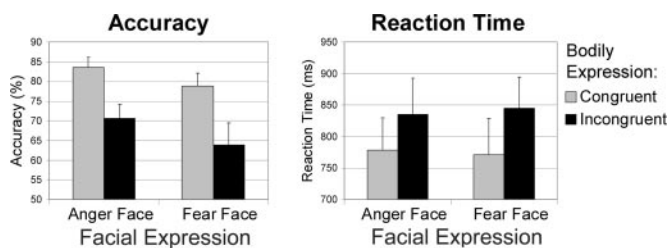

Fig. 2. Behavioral results of the facial expression task for the Face–Body Compound Stimuli. Participants had to judge the expression of faces that were accompanied by either a congruent or incongruent bodily expression.

Table 2. Behavioral results, isolated face and body stimuli

Stimulus type[R]	Emotion	Accuracy, %	Reaction time, ms
Body	Anger	81.6 ± 4.0	922 ± 80*a
	Fear	77.0 ± 2.9	855 ± 64*a
Face	Anger	80.9 ± 4.6	783 ± 50*b
	Fear	70.6 ± 4.3	798 ± 46*b

See Table 1 legend.

tained irrespective of whether the face was expressing anger or fear.

Table 2 shows the performance on the control stimuli. The emotional expression conveyed by the isolated faces and bodies was correctly judged in 78% of cases overall, without significant differences between conditions. The RT was faster for faces (790 ms) than for bodies [889 ms, $F(1, 11) = 8.68$, $P < 0.02$].

Electrophysiological Data. *P1 component.* The ERPs of 10 of 12 participants showed a distinctive prominent positive deflection in the occipital channels, peaking between 89 and 133 ms (average 115 ms). Fig. 3 shows a typical P1 response to face–body compound stimuli in an individual subject and the group average ERPs for both the compound and isolated stimuli at electrode sites O1, Oz, and O2. The quantitative data for peak amplitude
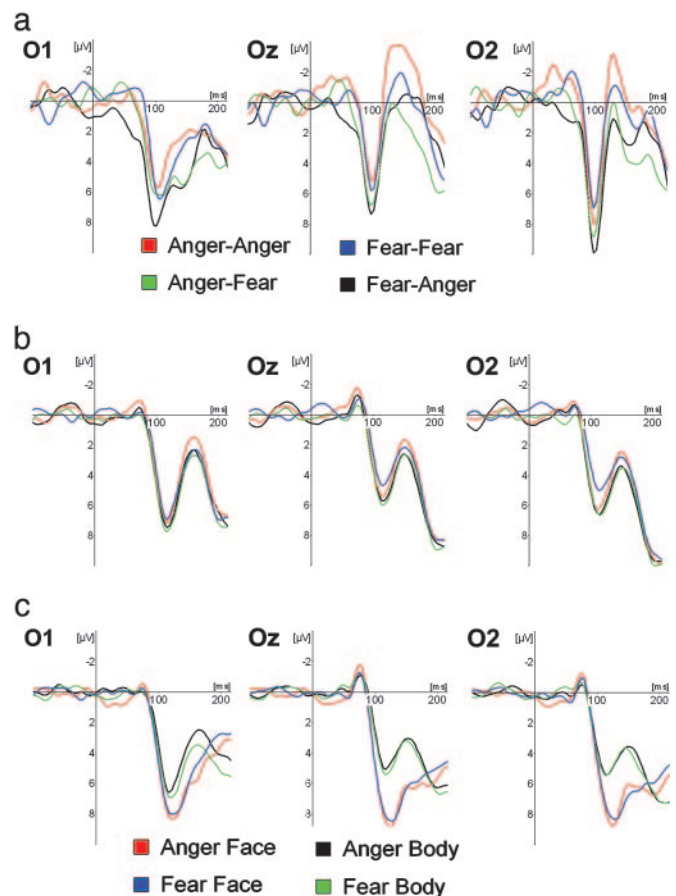

Fig. 3. P1 component at occipital electrodes O1, Oz, and O2. (*a*) Face–body compound stimuli, single subject. ERP signals of a single subject for Face–Body Compound stimuli showing a sharp positive deflection peaking around 100 ms. (*b*) Face–body compound stimuli, group average. Group average ERPs for compound stimuli. Legends as in top row. (*c*) Isolated faces and bodies, group average. Group average ERPs for Isolated Face and Body stimuli. Negativity is up.

Table 3. P1 component, face–body compound stimuli

EmoFace	EmoBody	Congruence[A]	Amplitude, μV	Latency, ms
Anger	Anger	Congruent	7.30 ± 0.91*[a]	115.2 ± 2.7
	Fear	Incongruent	7.72 ± 0.81*[b]	115.2 ± 2.7
Fear	Fear	Congruent	6.51 ± 0.64*[a]	116.8 ± 3.3
	Anger	Incongruent	7.65 ± 0.64*[b]	116.1 ± 3.1

Mean ± SEM values for amplitude and latency of the P1 component measured at electrodes O1, Oz, and O2, and results of the GLM: A and L in column heading denote main effects for amplitude and latency, respectively, for a given factor. Asterisks denote corresponding P values (*, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$), with symbols a and b, indicating the contrasting conditions. Group data of 10 participants are shown.

and latency across electrodes can be found in Tables 3 and 4. The following effects were found for the P1 amplitude. There was a lead main effect [$F (2, 8) = 4.92; P = 0.04$], with O1>O2>Oz, but post hoc tests failed to reveal significant differences between the electrodes, and no interaction effects were found for leads. There was a Congruence main effect [$F (1, 9) = 9.36, P = 0.014$] with larger P1 amplitudes for the incongruent than for the congruent compound stimuli. The amplitude appeared not to be sensitive to the facial or bodily expression. No latency effects were observed for the compound stimuli. The P1 component of the isolated face and body stimuli depended solely on StimulusType, with faces eliciting a larger [$F (1, 9) = 30.39; P < 0.001$] but slower [120 ms vs. 116 ms; $F (1, 9) = 10.31, P = 0.011$] response than bodies. An interaction effect for Lead×StimulusType was found for amplitude [$F (2, 8) = 11.87, P = 0.004$]; the stimulus effect was larger on Oz than on O1. The P1 component was not affected by the emotion expressed in the isolated faces and bodies.

N170 component. A distinctive N170 component with its typical lateral occipito-temporal distribution could be reliably identified in 11 of 12 subjects. Fig. 4 shows the group average ERPs for the compound and control stimuli at electrode sites P7 and P8. The quantitative data for peak amplitude and latency can be found in Tables 5 and 6. No effects were found for the amplitude of the N170, neither for the compound nor the control stimuli. The following effects were found for latency: A significant EmoBody main effect (EmoFace×Congruence interaction) was found for compound stimuli [$F (1, 10) = 22.36, P < 0.001$], with shorter latencies for faces combined with angry bodies (175 ms) than for faces combined with fearful bodies (180 ms). When faces and bodies were presented in isolation, however, the emotional expression did not have any influence on the latency. The timing appeared to depend solely on StimulusType [$F (1, 10) = 15.87; P = 0.003$], with bodies (187 ms) eliciting a faster response than faces (199 ms).

Discussion

As indicated by our behavioral results, congruent emotional body language improves recognition of facial expression, and conflicting emotional body language biases facial judgment toward the emotion conveyed by the body. Electrophysiological correlates provide evidence that this integration of affective information already takes place at the very earliest stages of face

Table 4. P1 component, isolated face and body stimuli

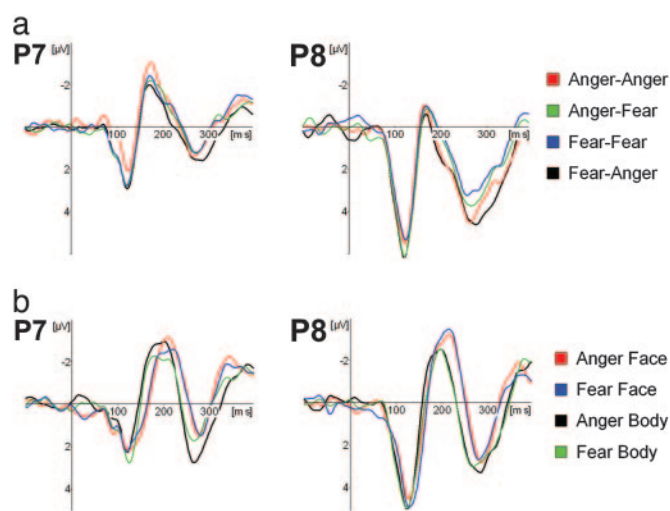| Stimulus type[A,L]	Emotion	Amplitude, μV	Latency, ms
Body	Anger	6.58 ± 0.81***[a]	116.1 ± 2.7*[a]
	Fear	6.80 ± 0.84***[a]	116.3 ± 2.8*[a]
Face	Anger	9.38 ± 1.20***[b]	119.4 ± 2.9*[b]
	Fear	8.99 ± 1.03***[b]	120.0 ± 3.0*[b]

See Table 3 legend.

Fig. 4. N170 component at occipito-temporal electrode sites P7 (left hemisphere) and P8 (right hemisphere). (*a*) Face–body compound stimuli. Group average ERPs for Face–Body Compound stimuli. (*b*) Isolated faces and bodies. Group average ERPs for Isolated Face and Body stimuli. Negativity is up.

processing. We found an overall congruence effect for the occipital P1 components at a latency of 115 ms, with larger amplitudes for the incongruent face–body composites as compared with congruent compound stimuli. A control ERP experiment showed this was not caused by differences in physical properties of the stimuli or the emotion expressed in isolated faces or bodies. In contrast, the N170 component appeared not to be sensitive for conflicting face–body emotions.

Incongruent Bodily Expression Hampers Facial Expression Recognition. Recognition of the emotion conveyed by the face is systematically influenced by the emotion expressed by the body. When observers have to make judgments about a facial expression, their perception is biased toward the emotional expression conveyed by the body. This effect has not previously been studied systematically in humans. It has been described in animal research, although that information from body expressions can play a role in reducing the ambiguity of facial expression (23). Moreover, it has been shown that observers' judgments of infant emotional states depend more on viewing the whole body than just the facial expressions (24). A similar effect on recognition of facial expressions was previously observed for the combination of a facial expression and an emotional tone of voice (25, 26).

The use of a forced-choice paradigm and a relatively short presentation (200 ms) suggests that the effect does not follow from an elaborate cognitive analysis of individual facial and bodily features but is based on fast global processing of the

Table 5. N170 component, face–body compound stimuli

EmoFace	EmoBody[L]	Congruence	Amplitude, μV	Latency, ms
Anger	Anger	Congruent	−3.94 ± 0.54	175.8 ± 6.3***[a]
	Fear	Incongruent	−3.57 ± 0.48	179.0 ± 5.8***[b]
Fear	Fear	Congruent	−3.78 ± 0.58	180.6 ± 7.4***[b]
	Anger	Incongruent	−3.19 ± 0.49	174.3 ± 6.1***[a]

Mean ± SEM values for amplitude and latency of the N170 component measured at electrodes P7 and P8 and results of the GLM: A and L in the column heading denote main effects for amplitude and latency, respectively, for a given factor. Asterisks indicate the corresponding P values (*, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$), with a and b indicating the contrasting conditions. Group data of 11 participants are shown.

Meeren *et al.*

PSYCHOLOGY

Table 6. N170 component, isolated face and body stimuli

Stimulus type[L]	Emotion	Amplitude, μV	Latency, ms
Body	Anger	-4.48 ± 0.53	185.9 ± 6.2**[a]
	Fear	-3.92 ± 0.48	188.2 ± 6.1**[a]
Face	Anger	-4.44 ± 0.58	199.2 ± 5.5**[b]
	Fear	-4.39 ± 0.49	199.2 ± 5.3**[b]

See Table 5 legend.

stimulus. That a reliable influence was obtained in an implicit paradigm in which the bodies were neither task-relevant nor explicitly attended to suggests the influence they exercise is rapid and automatic. Previous studies have shown that, even under unlimited viewing conditions, the immediate visual context of a "face" can induce powerful visual "illusions." For instance, the outer facial features appear to dominate our sense of facial identity at the cost of the inner ones (27–29), and showing bodies with faces hidden does activate the area in fusiform cortex related to face processing (19, 30).

Rapid Congruence Effect Indexed by an Enhanced P1 Component. The result of an overall congruence effect on the P1 component (enhanced for incongruent stimuli) with a latency of 116 ms (range 86–128 ms) points to a rapid automatic extraction of such vital information. It is unclear at present whether the observed effect is based on parallel extraction of the cues, provided respectively by face and body, or indicates that the face–body compound is processed as a whole. Indirectly, the pattern of results consisting of an increase in amplitude for the incongruent composites suggests processing of the whole rather than separate cue extraction. Indeed, the common interpretation of increase in amplitude is associated with the anomalous presentation condition. Assuming that the congruent face–body composites represent the normal condition, then deviation from this normal face–body composition in the incongruent cases will generate an increase in amplitude of the same waveform. Further evidence suggesting that the composites, rather than the separate face and body cues, are processed is provided by the finding that each of the cues on its own does not lead to the observed P1 effect.

In line with this, our findings are remarkable, because the P1 did not discriminate between the emotions anger and fear per se in the control stimuli (isolated faces and bodies). The latter is consistent with previous reports in which P1-amplitude modulation by emotion always refers to differences between emotional and neutral stimuli, not to differentiations between different emotions. Similar to our present results on intramodal integration, intermodal binding between affective information from the face and the voice has been shown to start as early as 110 ms poststimulus, because a specific enhancement in amplitude of the auditory N1 component was found (31).

Typically, the P1 amplitude increases when attention is directed to the location but not to nonspatial features of a stimulus (e.g., ref. 11). Behavioral studies in normal subjects (e.g., ref. 32) and brain-damaged patients with spatial neglect (33) suggest that emotional stimuli capture attention more readily than neutral stimuli. Images of emotional body language have proven to be very similar to images of facial expressions predominantly used in research on exogenous attention effects. It has recently been shown that human bodies also capture awareness in normal observers (34), and that fearful bodies reduce extinction in neglect patients (M. Tamietto, G. Geminiani, R. Genero, and B.d.G., unpublished results). Although incongruent stimuli did not receive more endogenous attention in our paradigm than congruent stimuli, a conflict between facial and bodily expression may attract more attention than emotion-congruent stimuli, thereby prompting more elaborate analysis in early visual areas.

Although there is no clear consensus about the exact location of the underlying neuronal generators, studies using source analysis suggest that the P1 originates mainly from "early" extrastriate visual areas, i.e., the lateral and dorsal (posterior) aspects of the occipital cortex (e.g., refs. 7–10). It is commonly assumed that these areas process only low-level features of the stimulus (i.e., luminance, contrast, line orientation, color, and movement). Indeed, early visual responses are very sensitive to changes in low-level properties. In accordance, we found a large difference between the P1 components of isolated faces and isolated bodies. Importantly however, the congruent and incongruent stimuli differed not in their low-level properties but only in their emotional content, because they consisted of exactly the same stimulus material.

Consistent with this, evidence is now accumulating indicating that early visual areas can also perform higher-level processing and extract meaningful information. Studies in monkeys have shown that neuronal responses in V1 and V2 are sensitive for contextual information (i.e., information outside the classical receptive field) and not only respond to the physical properties of the stimulus but also are correlated with its percept and "meaning" (e.g., refs. 35 and 36). Electrophysiological studies in humans have shown that the occipital component ≈ 100–120 ms already shows face selectivity (7, 12) and correlates with the conscious detection of a face. Furthermore, in addition to the "classic" N170 inversion effect, electrophysiological studies have observed an even earlier inversion effect for faces, i.e., an enhanced and delayed P1 component (13, 14, 37). This suggests that global face-specific processing already occurs ≈ 100 ms.

Further evidence for rapid visual processing of higher-order information in humans is provided by findings that average ERP responses to complex natural scenes already reflect the visual category of the stimulus shortly after visual processing has begun (e.g., 75–80 ms), even if this difference does not become correlated with the observer's behavior until 150 ms poststimulus (38).

N170: Lack of Emotional Congruence Effect in Structural Encoding Stage. The N170 component for the compound stimuli appeared to be sensitive neither for a mismatch between facial and bodily expressions nor for facial expression alone. Although a significant EmoBody main effect was found, its relevance should not be overestimated, because the time difference was only 5 ms, equaling one digital time sample. When isolated stimuli were presented, the N170 peaked significantly earlier for bodies than faces by 12 ms, a finding consistent with earlier work (6). The lack of a differential response between the two negative emotions anger and fear is in agreement with a large body of literature. Although one study has reported differential effects for the perception of fearful and angry faces (3), commonly the N170 is not found to be sensitive for emotional expression.

Integrative View and Concluding Remarks. Our behavioral and electrophysiological results suggest that when observers view a face in a natural body context, a rapid (<120 ms) automatic evaluation takes place whether the affective information conveyed by face and body are in agreement with each other. This early "categorization" into congruent and incongruent face–body compounds requires fast visual processing of the emotion expressed by face and body and the rapid integration of meaningful information. The time difference between the present P1 component and the first arrival of the initial afferent volley in V1 in humans measured from onset to onset (≈ 50 ms) or peak to peak (≈ 75 ms; refs. 8 and 39) is only 20–40 ms. Although activation spreads rapidly to other areas (39, 40), and processing in the visual system appears to be faster than traditionally assumed (e.g., refs. 38 and 41), this presumably does not allow enough time for the activity to pass through the various stages

of the ventral stream for object recognition, given the relatively slow passage time of each cortical stage. For comparison, face-selective neurons in monkeys become active only ≈40–60 ms after the initial activation of V1 (42, 43).

The fast extraction of biologically relevant information, however, is reconcilable with the dual route model for face (and body) recognition (20, 44) postulating that, apart from the face recognition route in temporal cortex, there exists a separate but interrelated face processing system that mediates the detection of faces and possibly of emotional bodily expressions. In contrast to the former, which finds its base in the relatively slow object recognition system, the face detection system relies on coarse processing, which allows for the analysis of facial expressions (45) and can therefore act fast. It functions optimally when stimuli are briefly flashed and capture attention. In this fast detection system, a crucial role has been hypothesized for the superior colliculus and the pulvinar. Viewing fearful faces and bodies activates these subcortical nuclei (20, 45, 46), and the unique connectivity of the pulvinar makes it highly suitable to signal the visual saliency of a stimulus and modulate the activity of extrastriate visual areas. In this context, there may be a special role for an extrageniculostriate pathway to the cortex that bypasses V1, as suggested by the residual ability of blindsight patient to process emotional stimuli (47, 48).

In conclusion, our data suggest there exists a neural mechanism for rapid automatic perceptual integration of high-level visual information of biological importance presented outside the focus of attention, thereby enabling evaluation of the relation between facial and bodily expressions to take place before full structural encoding of the stimulus and conscious awareness of the emotional expression is established.

1. Darwin, C. (1872/1965) *The Expression of the Emotions in Man and Animals* (Univ. of Chicago, Chicago).
2. Whalen, P. J., Shin, L. M., McInerney, S.C., Fischer, H., Wright C. I. & Rauch, S.L. (2001) *Emotion* **1,** 70–83.
3. Batty, M. & Taylor, M. J. (2003) *Brain Res. Cogn. Brain Res.* **17,** 613–620.
4. Bentin, S., Allison, T., Puce, A., Perez, E. & McCarthy, G. (1996) *J. Cognit. Neurosci.* **8,** 551–565.
5. Watanabe, S., Kakigi, R. & Puce, A. (2003) *Neuroscience* **116,** 879–895.
6. Stekelenburg, J. J. & de Gelder, B. (2004) *NeuroReport* **15,** 777–780.
7. Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J. & Ilmoniemi, R. J. (1998) *Neurosci. Lett.* **253,** 147–150.
8. Di Russo, F., Martinez, A., Sereno, M. I., Pitzalis, S. & Hillyard, S. A. (2001) *Hum. Brain Mapp.* **15,** 95–111.
9. Di Russo, F., Pitzalis, S., Spitoni, G., Aprile, T., Patria, F., Spinelli, D. & Hillyard, S. A. (2005) *NeuroImage* **24,** 874–886.
10. Martinez, A., Di Russo, F., Anllo-Vento, L. & Hillyard, S. A. (2001) *Clin. Neurophysiol.* **112,** 1980–1998.
11. Hillyard, S. A. & Anllo-Vento, L. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 781–787.
12. Liu, J., Harris, A. & Kanwisher, N. (2002) *Nat. Neurosci.* **5,** 910–916.
13. Itier, R. J. & Taylor, M. J. (2004) *NeuroImage* **21,** 1518–1532.
14. Itier, R. J. & Taylor, M. J. (2002) *NeuroImage* **15,** 352–372.
15. Pizzagalli, D. A., Lehmann, D., Hendrick, A. M., Regard, M., Pascual-Marqui, R. D. & Davidson, R. J. (2002) *NeuroImage* **16,** 663–677.
16. Halgren, E., Raij, T., Marinkovic, K., Jousmaeki, V. & Hari, R. (2000) *Cereb. Cortex* **10,** 69–81.
17. Eger, E., Jedynak, A., Iwaki, T. & Skrandies, W. (2003) *Neuropsychologia* **41,** 808–817.
18. Esslen, M., Pascual-Marqui, R. D., Hell, D., Kochi, K. & Lehmann, D. (2004) *NeuroImage* **21,** 1189–1203.
19. Hadjikhani, N. & de Gelder, B. (2003) *Curr. Biol.* **13,** 2201–2205.
20. de Gelder, B., Snyder, J., Greve, D., Gerard, G. & Hadjikhani, N. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 16701–16706.
21. Ekman, P. & Friesen, W. V. (1976) *Pictures of Facial Affect* (Consulting Psychologist, Palo Alto, CA).
22. Gratton, G., Coles, M. G. H. & Donchin, E. (1983) *Electroenceph. Clin. Neurophysiol.* **55,** 468–484.
23. Preston, S. D. & de Waal, F. B. (2002) *Behav. Brain Sci.* **25,** 1–20, discussion 20–71.
24. Camras, L. A., Meng, Z., Ujiie, T., Dharamsi, S., Miyake, K., Oster, H., Wang, L., Cruz, J., Murdoch, A. & Campos, J. (2002) *Emotion* **2,** 179–193.
25. Massaro, D. W. & Egan, P. B. (1996) *Psych. Bull. Rev.* **3,** 215–221.
26. de Gelder, B. & Vroomen, J. (2000). *Cognit. Emotion* **14,** 289–311.
27. Haig, N.D. (1986) *Perception* **15,** 235–247.
28. Sinha, P. & Poggio, T. (1996) *Nature* **384,** 404.
29. Sinha, P. (2002) *Nat. Neurosci. Suppl.* **5,** 1093–1097.
30. Cox, D., Meyers, E. & Sinha, P. (2004) *Science* **304,** 115–117.
31. Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B. & Crommelinck, M. (2000) *NeuroReport* **11,** 1329–1333.
32. Bradley, B. P., Mogg, K., Millar, N., Bonhamcarter, C., Fergusson, E., Jenkins, J. & Parr, M. (1997) *Cognit. Emotion* **11,** 25–42.
33. Vuilleumier, P. & Schwartz, S. (2001) *Neurology* **56,** 153–158.
34. Downing, P. E., Bray, D., Rogers, J. & Childs, C. (2004) *Cognition* **93,** B27–B38.
35. Albright, T. D. & Stoner, G. R. (2002) *Annu. Rev. Neurosci.* **25,** 339–379.
36. Bradley, D. (2001) *Curr. Biol.* **11,** R95–R98.
37. Itier, R. J. & Taylor, M. J. (2004) *J. Cognit. Neurosci.* **16,** 487–502.
38. VanRullen, R. & Thorpe, S. J. (2001) *J. Cognit. Neurosci.* **13,** 454–461.
39. Foxe, J. J. & Simpson, G. V. (2001) *Exp. Brain Res.* **142,** 139–150.
40. Vanni, S., Warnking, J., Dojat, M., Delon-Martin, C., Bullier, J. & Segebarth, C. (2004) *NeuroImage* **21,** 801–817.
41. Thorpe, S., Fize, D. & Marlot, C. (1996) *Nature* **381,** 520–522.
42. Bruce, C., Desimone, R. & Gross, C. G. (1981) *J. Neurophysiol.* **46,** 369–384.
43. Perrett, D. I., Rolls, E. T. & Caan, W. (1982) *Exp. Brain Res.* **47,** 329–342.
44. de Gelder, B. & Rouw, R. (2001) *Acta Psychol.* **107,** 183–207.
45. Vuilleumier, P., Armony, J. L., Driver, J. & Dolan, R. J. (2003) *Nat. Neurosci.* **6,** 624–631.
46. de Gelder, B., Vroomen, J., Pourtois, G. & Weiskrantz, L. (1999) *NeuroReport* **10,** 3759–3763.
47. Panksepp, J. (1998) *Affective Neuroscience: The Foundations of Human and Animal Emotions* (Oxford Univ. Press, New York).
48. Morris, J. S., de Gelder, B., Weiskrantz, L. & Dolan R. J. (2001) *Brain* **24,** 1241–1252.

PSYCHOLOGY