Tracing the time course of emotion perception:

The impact of stimulus physics and semantics on gesture processing

Tobias Flaisch & Harald T. Schupp

Department of Psychology, University of Konstanz; P.O. Box 36, 78457 Konstanz, Germany

Corresponding author: Tobias Flaisch
Department of Psychology
University of Konstanz
PO Box 36
78457 Konstanz, Germany
Phone: 0049 – 7531 – 88 – 5129
Fax: 0049 – 7531 – 88 – 2971
Email: tobias.flaisch@uni-konstanz.de

Key words: Attention, Emotion, ERP, Gestures, non-verbal communication
Numerous ERP-studies reveal the differential processing of emotional and neutral stimuli. Yet, it is an ongoing debate to what extent the ERP-components found in previous research are sensitive to physical stimulus characteristics or emotional meaning. The present study manipulated emotional meaning and stimulus orientation to disentangle the impact of stimulus physics and semantics on emotional stimulus processing. Negative communicative hand gestures of Insult were contrasted with neutral control gestures of Allusion to manipulate emotional meaning. An elementary physical manipulation of visual processing was implemented by presenting these stimuli vertically and horizontally. Results showed dissociable effects of stimulus meaning and orientation on the sequence of ERP-components. Effects of orientation were pronounced in the P1- and N170-time frames and attenuated during later stages. Emotional meaning affected the P1, evincing a distinct topography to orientation effects. While the N170 was not modulated by emotional meaning, the EPN- and LPP-components replicated previous findings with larger potentials elicited by the Insult gestures. These data suggest that the brain processes different attributes of an emotional picture in parallel and that a coarse semantic appreciation may already occur during relatively early stages of emotion perception.
**Introduction**

Reacting rapidly and appropriately to threatening or life-sustaining stimuli is imperative for an organism’s survival in a complex environment. Thus, motivationally relevant external stimuli may automatically induce a state of natural selective attention (Lang et al., 1997) and effectively guide ensuing perceptual processes. Research utilizing event-related brain potentials (ERP) has determined several indicators of preferential emotion processing along the time scale of visual perception (Flaisch et al., 2008b; Schupp et al., 2006). As emotionally guided perceptual amplification supposedly occurs from very early to relatively late stages of visual processing, the extent to which the observed emotional effects are driven by stimulus differences on a purely physical level or the semantic information associated with such differences is under debate. In the present study, we provide evidence that early indicators of emotion perception conceivably reflect semantic stimulus content and demonstrate that the effects of physical and semantic stimulus attributes can be differentiated along the sequence of emotion perception.

Preferential processing of affective pictorial stimuli unfolds in a temporal sequence from relatively early to later stages (Schupp et al., 2006). Studies utilizing ERP-measures identified several electro-cortical components presumably reflecting distinct sub-processes of emotion perception. A large body of research has established the ‘early posterior negativity’ (EPN) and the ‘late positive potential’ (LPP) as indices of emotion processing. The EPN is characterized by a more negative-going deflection of the surface potential when viewing emotional compared to neutral pictures (Flaisch et al., 2008a; Schupp et al., 2007). It is typically observed over temporo-parieto-occipital
sensors and occurs around 150-350 ms after stimulus onset. Subsequently, this difference is followed by the LPP which occurs around 350-750 ms over fronto-parieto-central positions as a relative positivity to emotional stimuli (Flaisch et al., 2008b; Schupp et al., 2006). These effects may reflect the amplified or prioritized processing of emotional stimuli as they are observed with striking similarity in passive task contexts across a wide variety of visual stimulus materials, including naturalistic scenes (IAPS, Flaisch et al., 2008b; Schupp et al., 2007), facial expressions (Leppänen et al., 2007; Schupp et al., 2004b), written words (Herbert et al., 2008; Kissler et al., 2009) and communicative hand gestures (Flaisch et al., 2009, 2011).

There is some evidence that emotional pictures are detected even earlier in the processing stream. Several studies have reported differences between emotional and neutral picture stimuli as early the P1, occurring around 60-120 ms over parieto-occipital sensor positions (Flaisch et al., 2011; Pourtois et al., 2005; van Heijnsbergen et al., 2007). In addition, several studies have investigated the N170-component observed over occipito-temporal sites with a latency around 170 ms, which is held to reflect the perceptual encoding of faces and body parts (for a recent review see Minnebusch and Daum, 2009). Mixed evidence has been obtained regarding the susceptibility of the N170 to the affective nature of face stimuli. While some studies do report emotional modulations of the N170 (Frühholz et al., 2011; Mühlberger et al., 2009; Stekelenburg and De Gelder, 2004), others have failed to find such effects (Leppänen et al., 2007; Pourtois et al., 2005; Schupp et al., 2004b). In sum, the perceptual processing of emotional picture stimuli unfolds in a temporal sequence which may be traced by various emotion-sensitive ERP-components.
However, such modulations may not exclusively indicate semantic-evaluative processes. Instead, similar ERP-effects may be observed in purely physical manipulations that either affect perceptual processes or interact with the semantic evaluation of according stimuli. Several studies hint at various physical parameters that may influence emotion perception at various stages, effectively impeding clear emotion interpretation. One such parameter is visual noise since degrading the perceptual quality of a picture stimulus generally diminishes the P3/LPP, a component widely held to reflect semantic stimulus encoding (Kok, 2001). Perceptually degraded pictures also invoke reduced or even abolished P1- and EPN-amplitudes (Schupp et al., 2008), most likely reflecting hampered object recognition. A differentiated picture emerges when image size is varied: Smaller pictures elicit less preferential processing at the level of the EPN, but later processes as indexed by the LPP seem to be largely unaffected by the physical size of the stimulus (De Cesarei and Codispoti, 2006). This suggests that these components reflect physical-perceptual and semantic processing differentially. Interestingly, emotional pictures presented in the periphery, compared to the fovea, appear to lose the ability to induce preferential processing (De Cesarei et al., 2009). Other studies have focused on the relation of spatial frequency and preferential emotion processing (Holmes et al., 2005; Pourtois et al., 2005; Vuilleumier et al., 2003). Results suggest that, in particular, early brain responses to emotional pictures rely on the information contained in the low-frequency spectrum of the images (Alorda et al., 2007; Pourtois et al., 2005; Vlamings et al., 2009). In contrast, subsequent ERP-indices, i.e., the N170 and the LPP, showed a much smaller or even no preference for frequency filtered emotional pictures (Alorda et al., 2007; Pourtois et al., 2005; Vlamings et al., 2009; De Cesarei et al., 2011). Picture composition may also obscure semantic emotion
effects. Comparing pictures with a simple figure-ground composition to more complex images revealed pronounced ERP-effects in the EPN-time window, while emotional effects were limited to the LPP (Bradley et al., 2007). This finding was interpreted as evidence that the EPN would mainly reflect the perceptual organization of a stimulus while the LPP would indicate its actual emotional appreciation. In sum, existing research shows that the different ERP-indices of preferential emotion processing may be susceptible to various physical stimulus manipulations. Thus, the exact nature of the various emotion-sensitive ERP-components, in particular whether they reflect purely physical-perceptual or semantic-evaluative processes, remains to be determined.

To fully appreciate physical and semantic processing differences along the sequence of emotion perception requires assessing effects from relatively early up to the later stages in the same participants utilizing the same stimulus materials. While previous studies sporadically speak to this issue, a comprehensive assessment of this question is still amiss. One reason for this may be found in the specifics of conventional stimulus materials. Results from research using emotional facial expressions are difficult to generalize as faces have a very unique status in the brain (Bentin et al., 1996; Farah et al., 1998; Kanwisher et al., 1997), presumably due to their evolutionary significance (Öhman et al., 2000). This restriction also applies to naturalistic pictures, most notably to those most reliably eliciting emotion processing, i.e. erotic or violent content (Schupp et al., 2004a). Furthermore, the vastly differing physical composition of IAPS pictures makes it difficult to disentangle physical from semantic effects based on these materials. Finally, no previous study has succeeded in delineating physical from semantic processes and thereby assessed the full sequence of emotion-modulated ERP-components.
Here, we approached this question by utilizing communicative hand gestures (Efron, 1972/1941; Morris, 1994). Specifically, we compared negative and neutral gestures rotated to either vertical or horizontal orientation. These materials appear particularly suited regarding the present research question. First, emotional gestures are processed preferentially by the brain (Flaisch et al., 2009), as demonstrated for the P1, the subsequent EPN, and the later LPP (Flaisch et al., 2011). Additionally, body stimuli elicit an N170 (Minnebusch and Daum, 2009), however, no emotion modulation has yet been reported for non-facial body stimuli. Second, the employed hand gestures share a very comparable and simple picture composition, minimizing the effects of uncontrolled stimulus physics, such as perceptual quality, visual eccentricity, spatial frequency, or visual complexity. Third, the semantic meaning attached to the employed hand gestures does not rely on evolutionary significance as it is based on cultural convention and social learning (Buck and VanLear, 2002). Together, this allowed manipulating stimulus orientation, a low-level visual attribute known to elicit early processing differences between horizontal and vertical stimuli (Grill-Spector and Malach, 2004; Kenemans et al., 1993, 1995) while at the same time preserving overall recognizability and semantic meaning. Based on previous research, emotion and stimulus orientation effects were expected to differ across the processing stream. While stimulus orientation effects were expected to be most pronounced during early processing stages, emotion effects are usually most prominent during subsequent stimulus processing. High-density ERP-recordings served to differentiate between emotion and orientation effects regarding time course and topographical appearance.
Materials and Methods

Participants. Twenty (10 females) volunteers aged between 20 and 32 years (M = 23.6) participated in the study. The experimental procedures were approved by the ethical committee of the University of Konstanz and all participants provided informed consent. They received monetary compensation or course credit for participation.

Stimuli. The middle finger jerk, which is amongst the strongest gestures of sexual insult and produced by the upward thrusting of the stiff middle finger served as the emotional gesture (“Insult”), while the index finger pointing in a specific direction served as the emotionally neutral control gesture (“Point”; Morris, 1979). Both gestures were rotated to appear in vertical and horizontal orientations (Fig. 1). Each gesture was posed by four women and four men. All gestures were displayed with the back of the hand facing the viewer and with a neutral, monotone gray-blue background. All pictures also appeared mirrored along the vertical axis to control for possible lateralization effects.

Self report. Following ERP measurement, participants were asked to rate the viewed gestures according to their perceived pleasantness and arousal using the Self-Assessment Manikin (SAM; Bradley and Lang, 1994). For statistical analysis, these ratings were entered into a two-factorial repeated measure ANOVA (Gesture: Point vs. Insult; Orientation: vertical vs. horizontal).

Procedure, ERP data acquisition and analysis. To limit eye movements to a single fixation, a stimulus was presented for 117 ms followed by a blank interstimulus-interval of 900 ms. The entire picture set consisting of the differentially oriented and mirrored gestures (8*2*2*2 = 64) was repeated 20 times resulting in a total of 1280
picture presentations. The pictures were shown in a randomized order in which no more than three pictures of the same experimental category were presented in succession and the transition frequencies between all categories were controlled. The session was divided into 4 blocks, allowing for posture adjustments during the pauses in between. Participants were instructed to simply view the pictures.

Brain and ocular scalp potential fields were measured with a 256-lead geodesic sensor net (GSN200v2.0; EGI: Electrical Geodesics, Inc., Eugene, OR), on-line bandpass filtered from 0.01 – 100 Hz, and sampled at 250 Hz using Netstation acquisition software and EGI amplifiers. Using EMEGS-software (Peyk et al., 2011), data editing and artifact rejection were based on an elaborate method for statistical artifact control, specifically tailored for the analysis of dense sensor ERP-recordings (Junghöfer et al., 2000). Data were converted to an average reference and stimulus synchronized epochs extracted from 200 ms pre- until 1000 ms post-stimulus onset. After baseline adjustment (100 ms pre-stimulus), separate average waveforms were calculated for all experimental categories for each sensor and participant.

In a first exploratory analysis, each time point and sensor was individually tested using a two-factorial ANOVA (Gesture: Point vs. Insult; Orientation: vertical vs. horizontal). Significant effects were thresholded at $p < .05$ for at least eight continuous data points (32 ms); two neighboring sensors provided a conservative guarding against chance findings (Sabbagh and Taylor, 2000). Based on the resulting $F$-Values, the Global Power across all sensors was calculated to determine the time course of modulation as a function of both the main effects and the interaction. From this, five modulations became readily apparent, two as a function of Orientation and three for...
Gesture, respectively (Fig. 3). To further detail these effects, specific time intervals and sensor clusters were defined. Averaged data within these intervals and clusters were subjected to three-factorial ANOVAs with repeated measurement on the factors Gesture, Orientation, and Cluster.

Visual inspection of the two main effects of Gesture and Orientation in the P1-time frame suggested that the two effects differed regarding topographical appearance. Specifically, while the main effect of Gesture appeared to closely mirror the topographical distribution of the visual P1, the main effect of Orientation suggested a more superior and comparatively central localization (Fig. 5). Thus, to accurately capture possible distinctions between the two effects, they were assessed in three sensor clusters in the interval 100-120 ms. Two lateralized parieto-occipital clusters were centered on the P1-peak of the overall ERP across all conditions (Fig. 4; left: EGI # 106, 107, 108, 116, 117, 118, 125, 126; right: EGI # 139, 140, 150, 151, 152, 160, 161, 170), and a centro-parietal sensor cluster was positioned at sensor sites showing the most pronounced effects of stimulus orientation (EGI # 98, 99, 100, 110, 111, 120, 129, 130, 142).

The N170 was scored in the interval 160-200 ms in two lateralized occipito-temporal clusters (Fig. 4; left: EGI # 105, 106, 107, 112, 113, 114, 115, 116, 117, 121, 122, 123, 124, 125, 126, 134, 135, 136, 137, 146, 147; right: EGI # 139, 149, 150, 151, 157, 158, 159, 160, 161, 166, 167, 168, 169, 170, 175, 176, 177, 178, 188, 189, 200).

The EPN was captured in the interval 220-320 ms in two lateralized parieto-occipital clusters (Fig. 4; left: EGI # 95, 96, 97, 98, 106, 107, 108, 109, 110, 115, 116,
The LPP was scored in the interval 460-560 ms in two lateralized fronto-central sensor clusters (Fig. 4; left: EGI # 9, 16, 17, 24, 41, 42, 43, 44, 49, 50, 51, 52, 57, 58, 59, 79; right: EGI # 7, 132, 133, 145, 156, 184, 185, 186, 187, 196, 197, 198, 199, 206, 207, 208).

To assess topographic differences between ensuing effects in more detail, current source densities (CSD) were calculated (Junghöfer et al., 1997). This approach is based on a physiological volume conductor model that is well suited for dense array EEG data and which indicates a focal generator source by a sink/source pattern of inward/outward flow of current.

Control experiment

To provide an empirical control for orientation effects independent of semantic meaning, a second study utilizing simple horizontal and vertical bar stimuli was conducted (Fig. 1). The width and color of the bars closely mirrored the gesture images, and the same background was utilized. Twelve volunteers (7 female; age: 20–37 years, M = 25.7) participated in the study. Participants passively viewed 640 (2*320) picture presentations of the bar stimuli and all further experimental parameters were identical to the main study. Continuous EEG was recorded using 256-channel HydroCel-Geodesic Sensor Nets (EGI) and was processed as described for the main study.

Visual inspection of the difference wave [vertical–horizontal] revealed a modulation in the P1-time frame over parieto-central sensors, which preceded the P1-
peak (~110 ms). To statistically assess this effect and to test for its topographical accordance with the visual P1, the same approach was chosen as for the P1-analysis in the main study. Specifically, the time interval 72-92 ms post-stimulus was analyzed in two laterialized parieto-occipital sensor clusters, which were centered on the peak of the visual P1 (left parieto-occipital: EGI # 115, 116, 122, 123, 124, 125, 135, 136; right parieto-occipital: EGI # 138, 148, 149, 150, 157, 158, 159, 167) and in a centro-parietal sensor cluster which was positioned at sensor sites showing the most pronounced effects of stimulus orientation (EGI # 80, 81, 89, 90, 100, 101, 129, 130, 131). The extracted data were entered into a two-factorial ANOVA with repeated measurement on the factors Orientation and Cluster.

**Results**

**SAM-Ratings.** Participants perceived the Insult as more unpleasant than the Point gestures, regardless of stimulus orientation (Gesture: $F(1, 19) = 21.2, p < .001$; Point: $M = 4.63$, $SEM = 0.2$; Insult: $M = 3.23$, $SEM = 0.23$; Fig. 2). Similarly, the Insult gestures were also rated as more arousing (Gesture: $F(1, 19) = 21.9, p < .001$; Point: $M = 3.98$, $SEM = 0.43$; Insult: $M = 5.86$, $SEM = 0.32$; Fig. 2). Additionally, the vertical gestures were also perceived as slightly more arousing than the horizontal stimuli (Orientation: $F(1, 19) = 5.1, p < .05$; vertical: $M = 5.25$, $SEM = 0.35$; horizontal: $M = 4.59$, $SEM = 0.35$).

**Event-related potentials.**

**P1-Modulation.** A first modulation of the scalp potential emerged around 80-120 ms post-stimulus. Visual inspection of the Global Power (Fig. 3) suggested significant effects of both experimental factors in this time frame. However, the two main effects
appeared to differ regarding their scalp topography (Fig. 5C). While effects of gesture type closely mirrored the appearance of the visual P1 (Fig. 5), effects of stimulus orientation appeared maximally pronounced at more superiorly and centrally located parietal sensor positions. To capture these differences, the ERP was scored in three sensor clusters in the time window 100-120 ms and submitted to a three-factorial repeated-measures ANOVA (Gesture: Point vs. Insult; Orientation: vertical vs. horizontal; Cluster: left parieto-occipital vs. right parieto-occipital vs. centro-parietal). Interactions of Gesture by Cluster ($F(2, 38) = 5.9, p = .01$) and Orientation by Cluster ($F(2, 38) = 5.3, p = .01$) confirmed the observations suggested by visual inspection. These interactions were followed up by reduced-model ANOVAs for each of the three clusters, incorporating the factors Gesture and Orientation, respectively.

Main effects of Gesture were found in all three sensor clusters, albeit most pronounced over left and right parieto-occipital (P1-centered) sensor clusters (left: $F(1, 19) = 11.9, p < .01$; Point: $M = 1.99$, $SEM = 0.35$; Insult: $M = 2.32$, $SEM = 0.42$; right: $F(1, 19) = 22.2, p < .001$; Point: $M = 2.22$, $SEM = 0.28$; Insult: $M = 2.63$, $SEM = 0.34$; central: $F(1, 19) = 5.6, p < .05$; Point: $M = 0.94$, $SEM = 0.31$; Insult: $M = 1.15$, $SEM = 0.35$).

A somewhat different picture emerged for the main effect of Orientation, which was also significant in all sensor clusters but maximally pronounced for the parieto-central cluster (left: $F(1, 19) = 28.0, p < .001$; vertical: $M = 2.62$, $SEM = 0.42$; horizontal: $M = 1.68$, $SEM = 0.37$; right: $F(1, 19) = 16.7, p < .001$; vertical: $M = 2.78$, $SEM = 0.33$; horizontal: $M = 2.06$, $SEM = 0.31$; central: $F(1, 19) = 54.7, p < .001$; vertical: $M = 1.66$, $SEM = 0.33$; horizontal: $M = 0.43$, $SEM = 0.35$).
**N170-Modulation.** The next modulation emerged around 150-210 ms post-stimulus as a main effect of stimulus orientation (Fig. 3). To assess this modulation in detail, the N170 was scored in the time window 160-200 ms in two laterialized occipito-temporal sensor clusters, and the data were submitted to a three-factor repeated measures ANOVA (*Gesture*: Point vs. Insult; *Orientation*: vertical vs. horizontal; *Cluster*: left vs. right). Results revealed a highly significant main effect for *Orientation* ($F(1, 19) = 53.8, p < .001$; vertical: $M = 0.45, SEM = 0.43$; horizontal: $M = 0.99, SEM = 0.43$), indicating an increased negative-going deflection of the ERP to vertically oriented gestures (Fig. 6).

**EPN-Modulation.** Further down the temporal cascade, the Global Power yielded a third modulation over posterior leads (Figs. 3 and 7B), beginning around 200 ms and lasting until around 350 ms post-stimulus. This effect was captured in a time window from 220-320 ms in two laterialized parieto-occipital sensor clusters with a slightly more superior localization. Again, the data were submitted to a three-way ANOVA (*Gesture*: Point vs. Insult; *Orientation*: vertical vs. horizontal; *Cluster*: left vs. right). A highly significant main effect of *Gesture* ($F(1, 19) = 45.6, p < .001$; Point: $M = 2.15, SEM = 0.36$; Insult: $M = 1.62, SEM = 0.37$) indicated that the modulation was exclusively due to more negative amplitudes to Insult gestures (Fig. 7).

**LPP-Modulation.** A final modulation emerged around 440-580 ms and was located over fronto-central regions (Figs. 3 and 8B). Accordingly, the ERP was assessed in two laterialized fronto-central sensor clusters in the time window 460-560 ms. The three-way ANOVA (*Gesture*: Point vs. Insult; *Orientation*: vertical vs. horizontal; *Cluster*: left vs. right) revealed a significant higher-order interaction between *Gesture*,
Orientation, and Cluster \((F(1, 19) = 6.9, p < .05)\). This was followed up by reduced-model ANOVAs for each lateraled cluster, incorporating the factors Gesture and Orientation, respectively. Significant main effects of Gesture \((F's(1, 19) > 20.3, p's < .001)\) were observed for left (Point: \(M = -0.36, SEM = 0.18\); Insult: \(M = 0.02, SEM = 0.19\)) and right (Point: \(M = -0.09, SEM = 0.18\); Insult: \(M = 0.32, SEM = 0.21\)) sensor clusters, indicating an increased positivity for Insult as compared to Point gestures. In contrast, an effect of Orientation \((F(1, 19) = 7.3, p < .05)\); vertical: \(M = 0.2, SEM = 0.2\); horizontal: \(M = 0.03, SEM = 0.19\) was specific to right sensor sites, with vertical compared to horizontal gestures eliciting a slightly increased positive deflection.

**Control experiment**

Visual inspection revealed an early difference between vertical and horizontal bars around 60-100 ms post-stimulus and over centro-parietal leads, preceding the P1-peak by ~30-40 ms. To capture this difference and to test its topographical accordance with the visual P1, the ERP was scored in three sensor clusters in the time window 72-92 ms and submitted to a repeated-measures ANOVA (Orientation: vertical vs. horizontal; Cluster: left parieto-occipital vs. right parieto-occipital vs. centro-parietal). A significant interaction of Orientation by Cluster \((F(2, 22) = 7.3, p < .01)\) confirmed that vertical bars elicited increased amplitudes over parieto-central leads \((F(1, 11) = 10.6, p < .01)\); vertical: \(M = 0.14, SEM = 0.10\); horizontal: \(M = -0.27, SEM = 0.13\) but not over P1-centered left and right sensor clusters \((F's(1, 11) < 0.5, ns.; vertical left: M = -0.13, SEM = 0.18; horizontal left: M = -0.18, SEM = 0.23; vertical right: M = 0.00; SEM = 0.22, horizontal right: M = -0.11, SEM = 0.25)\). Inspection of the N170-, EPN-, and LPP-time frames revealed no further orientation differences.
Discussion

The present study aimed to dissociate the processing differences of emotional gestures according to simple physical features and effects due to the semantic content of the stimuli. Varying the orientation of emotional and neutral hand gestures allowed for the manipulation of a fundamental sensory attribute while largely preserving the inherent semantic meaning of the stimuli. Consistent with the notion of hierarchical visual processing (Grill-Spector and Malach, 2004), results demonstrate physical processing differences predominantly during early stages of visual perception, while semantic content gained increasing ascendancy over relatively later processing stages. Interestingly, the impact of semantic content was not exclusively restricted to later stages, commonly assumed to reflect elaborate stimulus processing, but was also distinctly apparent at the early P1- and EPN-components. These results provide evidence that the brain processes physical and semantic stimulus attributes in parallel, and that emotional meaning may be detected at relatively early stages of visual perception.

The first differentiation between the experimental categories emerged around the P1-peak (80-120 ms) over parieto-occipital sensors, appearing as relatively augmented amplitudes to vertical stimuli on the one hand, and Insult gestures on the other hand. However, while the modulation due to gesture type appeared to act directly on the perceptual processes reflected by the visual P1, effects of stimulus orientation appeared as differential ERP-activity superimposing the P1 (Luck, 2005). This interpretation is supported by the results of the control experiment, which also showed increased amplitudes in the P1-time frame to vertical stimuli over centro-parietal but not over P1-
centered parieto-occipital sensor clusters. In the control experiment, effects of stimulus orientation were also characterized by a relatively earlier onset. Interestingly, analyzing an even earlier time frame in the main study (i.e., 80-100 ms) revealed very comparable orientation effects but no hint of modulation due to gesture type. Thus, it appears likely that the orientation effect observed in the two experiments reflects purely physically driven low-level perceptual processing in early visual cortical areas. As such, this observation is reminiscent of studies comparing vertical to horizontal spatial gratings (Kenemans et al., 1993, 1995). Interestingly, another study found that mental rotation elicited increased neural activity in dorsal visual pathways located in the superior parietal lobe (Gauthier et al., 2002). In contrast, the P1-modulation, due to gesture type, may possibly reflect an early semantic differentiation of physically simple stimuli based on coarse stimulus features. While it may not conclusively be ruled out that minute physical differences other than stimulus orientation may account for the present findings, several studies speak to the brain’s ability to distinguish between stimuli at this early stage. Consistent with the assumption of motivationally guided attention allocation to Insult gestures, enhanced P1-amplitudes have been observed repeatedly in classical spatial attention research for attended as compared to unattended stimuli (for an overview see Luck et al., 2000). More importantly, it has been shown that the P1 is affected by object-based attention (Taylor, 2002) and a recent study reported that the P1 is selectively modulated by relatively complex visual stimuli depicting fearful, disgusting and neutral contents (Krusemark & Li, 2011). These observations are also highly consistent with other results showing emotionally augmented P1-amplitudes over lateral occipital regions (Pourtois et al., 2004, 2005). Finally, in the study by Gauthier and colleagues (2002), object recognition was associated with activations along ventral
visual areas but was not affected by mental rotation suggesting different mechanisms underlying the resolution of geometric deviance on the one hand, and object perception, on the other hand. In sum, the present effects in the P1-time frame are consistent with the notion of parallel and distinct processing mechanisms subserving mental rotation and visual orientation on the one hand, and early emotion recognition on the other hand.

Subsequently, a second ERP-modulation over temporo-occipital leads was revealed with a latency around 160-200 ms post-stimulus. This difference appeared as a more negative-going ERP-deflection to vertical as compared to horizontal stimuli. Time course and localization of this effect are reminiscent of the facial processing related N170. While the N170 is widely regarded as an indicator of face specific categorical processing (see e.g. Bentin et al., 1996; Itier and Taylor, 2004), recent studies have reported a very similar component related to the processing of bodies and body parts (De Gelder et al., 2010; Meeren et al., 2005; Minnebusch and Daum, 2009; Stekelenburg and De Gelder, 2004; Thierry et al., 2006). Interestingly, only hand gestures but not meaningless bars elicited a notable N170-difference of stimulus orientation suggesting the structural encoding of body parts. However, the N170-component was not affected by emotional meaning. While mixed results have been reported regarding the emotional modulation of the N170 both for facial, as well as for body part stimuli (De Gelder et al., 2010), the present results may relate to a recent study directly comparing emotional and neutral face and body expressions (Stekelenburg and De Gelder, 2004). Interestingly, while emotional modulations of the N170 to face stimuli were readily apparent in this study, no such effect was found for the body related N170. The finding that the N170 was strongly influenced by stimulus orientation may in turn relate to studies examining mental rotation of hand stimuli.
(Thayer et al., 2001; Thayer and Johnson, 2006). In these studies, increased ERP-negativities with very similar time course and topography to the present N170 were found when processing rotated hands. Interestingly, the same studies also reported rotation effects on the earlier P1. Taken together, the present N170-effects are consistent with the notion of physically driven processing differences during the encoding of body parts in higher-order visual regions.

A further ERP-difference emerged as an early posterior negativity for emotional as compared to neutral gestures around 200-360 ms post-stimulus. This effect closely resembles emotion modulations as reported in previous studies using gestures (Flaisch et al., 2009, 2011) but also physically and semantically different stimulus materials including naturalistic scenes (Flaisch et al., 2008a; Schupp et al., 2007), facial expressions (Leppänen et al., 2007; Schupp et al., 2004b), and written words (Herbert et al., 2008; Kissler et al., 2007, 2009). However, comparing images with simple figure-ground configuration to complex scenes, another study reported very comparable effects due to picture composition (Bradley et al., 2007). In this regard, the present EPN-findings provide new evidence supporting the notion of semantic differentiation at the level of the EPN. Utilizing stimuli with very simple basic physical configuration on the one hand, and very high perceptual similarity on the other hand, enables the effects of emotion selection to be dissociated from purely physically driven ERP differences. This notion is further strengthened by not only demonstrating a null effect of stimulus orientation on the EPN, but more so by showing that the applied physical manipulation indeed had a profound impact on the ERP, i.e. the P1- and the N170-components. This dissociates physical from semantic processing along the time course of stimulus perception. The present data are in this respect consistent with explicit attention studies.
showing that simple and complex target stimuli are differentiated as early as 150 ms post-stimulus (Smid et al., 1999; Thorpe et al., 1996; Codispoti et al., 2006). In sum, the present data suggest that the EPN, as observed for a variety of stimulus categories, likely reflects stimulus selection based on the emotional meaning of the stimuli.

The present results revealed the latest modulations in the time frame 440-580 ms post-stimulus as increased positivity over fronto-central sensors, both as a function of emotional meaning as well as stimulus orientation. The appearance of this effect is reminiscent of the LPP, a component typically found across a variety of visual stimulus materials when comparing emotional to neutral stimuli (Flaisch et al., 2008b; Herbert et al., 2008; Leppänen et al., 2007; Schupp et al., 2004b). The LPP has been related to elaborate stimulus processing and semantic evaluation (Schupp et al., 2006) and fits into the overall scheme of P300-like modulations, which assumedly index stimulus discrimination and resource allocation at a post-perceptual processing stage (Johnson, 1988; Schupp et al., 2006; Bradley et al., 2007). Thus, the present results may indicate the selection and increased allocation of attention to emotionally charged Insult gestures. While the effect of stimulus orientation also acting at this stage was comparatively small and distinctly observed with a right-lateralized topographical distribution, it may appear puzzling. However, the self-report data and the absence of any LPP-like modulation in the control experiment hint at the possibility of an emotional differentiation between even vertical and horizontal gestures. Specifically, vertical gestures were rated as more arousing than the horizontal ones, suggesting that vertical gestures may have a higher inherent significance, possibly reflected by the present orientation effect on the LPP. Consistent with this notion, signalling and warning hand signs intended to catch other people’s attention typically involve raising
the arm above one’s head with an upright gesture. Taken together, the observed LPP-effects are consistent with the notion of stimulus selection and the elaborate processing of emotionally significant visual stimuli.

The aims of the present study were twofold: i) isolating physical from semantic processing differences using a set of highly comparable emotional stimuli under preservation of their natural appearance, and ii) tracing potential effects of stimulus physics and semantics across the sequence of emotion perception as typically observed in passive viewing experimental contexts. Results show dissociable effects of stimulus meaning and orientation on all emotion-relevant ERP-components. The result pattern suggests a pre-dominance of low level physical stimulus features in early time windows encompassing the P1 and the N170, while the impact of stimulus meaning pronouncedly surpassed that of orientation on the later EPN- and LPP-components. Effects of stimulus semantics were observable as early as the P1 and most importantly for the EPN. These results suggest that the brain processes different attributes of an emotionally significant picture stimulus in parallel and that a coarse semantic appreciation may already occur during early stages of emotion perception.

References


Funding

This research was supported by the German Research Foundation (Schu 1074/10-3 and 1074/11-2).
Figure captions

Figure 1. Examples of gesture stimuli

Stimuli from one actor for all experimental cells in the main study (top) and utilized bar stimuli of the control experiment (bottom).

Figure 2. Subjective self-report

Mean valence and arousal ratings as a function of gesture type and stimulus orientation.

Figure 3. Global power of exploratory single sensor analysis

Global power plot of the $F$-values resulting from the 2x2-ANOVAs ($\text{Gesture} \times \text{Orientation}$) calculated for each sensor and time point separately. Statistical modulations for the main effect of stimulus orientation were clearly apparent in relatively early time frames, i.e. around 80-120 ms and 160-210 ms (black line), while modulations due to gesture type were most pronounced during later time windows, i.e. around 200-320ms and 450-600 ms (gray line). No clear interaction effects were obvious from this analysis.

Figure 4. Analyzed sensor clusters

Illustration of all sensor clusters as analyzed for each ERP-effect. Black sensors indicate channels presented in detail in Figs. 5-8. The sensor configuration is displayed from the top with the nose pointing up.

Figure 5. P1-component
A) Illustration of the ERP waveforms for a right parieto-occipital sensor (EGI #152, Fig. 2). B) Scalp topography maps (top) and according source analysis (bottom) for all picture categories in the time window 100-120 ms. C) Scalp difference maps (vertical–horizontal and Insult–Point; left) and according source analysis (right) in the time window 100-120 ms. Please note the different scales.

Figure 6. N170-component

A) Illustration of the ERP waveforms for a right temporo-occipital sensor (EGI #159, Fig. 2). B) Scalp difference maps (vertical–horizontal, left) and according source analysis (right) in the interval 160-200 ms.

Figure 7. EPN-component

A) Illustration of the ERP waveforms for a right temporo-occipital sensor (EGI #161, Fig. 2). B) Scalp difference maps (Insult-Point, left) and according source analysis (right) in the interval 220-320 ms.

Figure 8. LPP-component

A) Illustration of the ERP waveforms for a right fronto-central sensor (EGI #186, Fig. 2). B) Scalp difference maps (Insult-Point, left side; vertical-horizontal, right) in the interval 460-560 ms.
Stimuli from one actor for all experimental cells in the main study (top) and bar stimuli utilized in the control experiment (bottom).

42x53mm (300 x 300 DPI)
Mean valence and arousal ratings as a function of gesture type and stimulus orientation.
59x39mm (300 x 300 DPI)
Global power plot of the F-values resulting from the 2x2-ANOVAs (Gesture X Orientation) calculated for each sensor and time point separately. Statistical modulations for the main effect of stimulus orientation were clearly apparent in relatively early time frames, i.e. around 80-120 ms and 160-210 ms (black line), while modulations due to gesture type were most pronounced during later time windows, i.e. around 200-320ms and 450-600 ms (gray line). No clear interaction effects were obvious from this analysis.
Illustration of all sensor clusters as analyzed for each ERP-effect. Black sensors indicate channels presented in detail in Figs. 5-8. The sensor configuration is displayed from the top with the nose pointing up.

69x62mm (300 x 300 DPI)
A) Illustration of the ERP waveforms for a right parieto-occipital sensor (EGI #152, Fig. 2). B) Scalp topography maps (top) and according source analysis (bottom) for all picture categories in the time window 100-120 ms. C) Scalp difference maps (vertical–horizontal and Insult–Point; left) and according source analysis (right) in the time window 100-120 ms. Please note the different scales.
A) Illustration of the ERP waveforms for a right temporo-occipital sensor (EGI #159, Fig. 2). B) Scalp difference maps (vertical–horizontal, left) and according source analysis (right) in the interval 160-200 ms.

73x63mm (300 x 300 DPI)
A) Illustration of the ERP waveforms for a right temporo-occipital sensor (EGI #161, Fig. 2). B) Scalp difference maps (Insult-Point, left) and according source analysis (right) in the interval 220-320 ms.
A) Illustration of the ERP waveforms for a right fronto-central sensor (EGI #186, Fig. 2). B) Scalp difference maps (Insult-Point, left side; vertical-horizontal, right) in the interval 460-560 ms.

83x83mm (300 x 300 DPI)