Research report

Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis

Daniel Jokisch[a,b,*], Irene Daum[a,b], Boris Suchan[a], Nikolaus F. Troje[a,b,c]

[a] *Institute of Cognitive Neuroscience, Ruhr-University Bochum, Universitätsstr. 150, D-44780 Bochum, Germany*
[b] *International Graduate School of Neuroscience, Ruhr-University Bochum, Universitätsstr. 150, D-44780 Bochum, Germany*
[c] *Department of Psychology, Queen's University, Kingston, Ontario, Canada K7L 3N6*

Abstract

In the present study, we investigated how different processing stages involved in the perceptual analysis of biological motion (BM) are reflected by modulations in event-related potentials (ERP) in order to elucidate the time course and location of neural processing of BM. Data analysis was carried out using conventional averaging techniques as well as source localization with low resolution brain electromagnetic tomography (LORETA). ERPs were recorded in response to point-light displays of a walking person, an inverted walking person and displays of scrambled motion. Analysis yielded a pronounced negativity with a peak at 180 ms after stimulus onset which was more pronounced for upright walkers than for inverted walkers and scrambled motion. A later negative component between 230 and 360 ms after stimulus onset had a larger amplitude for upright and inverted walkers as compared to scrambled walkers. In the later component, negativity was more pronounced in the right hemisphere revealing asymmetries in BM perception. LORETA analysis yielded evidence for sources specific to BM within the right fusiform gyrus and the right superior temporal gyrus for the second component, whereas sources for BM in the early component were located in areas associated with attentional aspects of visual processing. The early component might reflect the pop-out effect of a moving dot pattern representing the highly familiar form of a human figure, whereas the later component might be associated with the specific analysis of motion patterns providing biologically relevant information.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Biological motion; Visual perception; Event-related potentials; LORETA; Recognition; Encoding; Inversion effect

1. Introduction

The human visual system is very sensitive to the detection of animate motion patterns. We can efficiently detect another living being in a visual scene, recognize human action patterns and attribute many features of psychological, biological and social relevance to other persons.

An experimental approach for studying information from biological motion (BM) with reduced interference from non dynamic cues is to represent the main joints of a person's body by bright dots against a dark background [32]. From such point-light displays, observers can easily recognize a human walker, determine his/her gender [5,16,33,34,58], recognize various action patterns [18], identify individual persons [17] and even recognize themselves [6]. The evolutionary importance of the perception of animate motion patterns has led to the development of a specific neural machinery as shown by several brain imaging studies using fMRI and PET [11,24–27,51,62]. These studies report selective activation of the superior temporal sulcus (STS) to visual stimuli consisting of BM. In addition to area STS, activation specific to BM has also been shown in the cerebellum [25,62], area VP [51], the amygdala [11] and the occipital and fusiform face area [27]. Activity in the STS-region is not induced by meaningful and coordinated non-biological motion such as the pendulum movements of a grandfather clock [44]. A dissociation between visual processing of

* Corresponding author. Tel.: +49 234 32 23174; fax: +49 234 32 14622.
E-mail address: daniel.jokisch@ruhr-uni-bochum.de (D. Jokisch).

moving humans and moving manipulable objects was also supported by Beauchamp et al. [7] by showing STS activity to human point-light and video displays in contrast to activity in the middle temporal gyrus to tool video and point-light displays.

According to a neural model based on BM perception studies, both the dorsal and ventral processing streams contribute to the perceptual analysis of BM [23]. The ventral form pathway is thought to provide information about sequences of body postures; the dorsal motion pathway is thought to provide information about complex optic flow patterns. Data from both pathways are integrated in the STS region.

This region is not only involved in the perception of whole body movements. Activation of STS is also observed during perception of movements of the eyes, hand and mouth and even when looking at implied motion in static images [2]. With respect to face perception, STS is suggested to be involved in the processing of dynamic aspects of faces that convey information facilitating social communication [28].

Results from imaging studies as well as from simulations [23] are consistent with neuropsychological findings in neurological patients suffering from focal brain lesions [15,35,50,60,61]. These case studies provide evidence for a dissociation between mechanisms involved in the perception of BM on the one hand, and mechanisms involved in inanimate visual motion tasks or static object recognition tasks on the other hand.

One approach to study the neuronal dynamics of perception of action is to measure activity when presenting stimuli consisting of body movements in full view in which the display initially stands still and movement onset occurs with a delay. Neural responses to onset of movements of the mouth and the eyes [47] were observed within 200 ms after motion onset as measured by ERPs. Facial movements occurring on a continuously present face elicited different N170 amplitudes for mouth opening versus closing and for eye aversion versus eyes gazing at the observer. Similar results were found for the observation of whole body actions of others [64]. ERPs elicited in response to movement onset in movie sequences of body stepping, hand closing and opening, and mouth opening and closing were selective for specific hand and body motions.

Findings from ERP and functional imaging studies in humans are complemented by electrophysiological studies in monkeys. Single cell recordings in the macaque superior temporal polysensory area (STP) yielded neurons which responded selectively to the sight of whole body movements as well as to point-light displays of BM [38]. Many STS cells integrate information about the form and motion of animate objects [39].

Further support for the notion that also the temporal lobe integrates facial form and motion in humans stems from an fMRI and ERP study of visual processing of natural and line-drawings displays of moving faces [48]. The STS and the fusiform gyrus responded selectively to both types of face stimuli, and they evoked larger ERPs compared to control stimuli at around 200 ms post motion onset. Puce and Perrett [46] recently concluded that specialized visual mechanism exist in the STS complex of both humans and non-human primates which produces selective neural responses to moving natural images of faces and bodies. These mechanisms are also involved in the processing of point-light displays of BM.

Whereas substantial knowledge is available about the neural structures underlying BM perception, many issues concerning the processing stages and their temporal characteristics in humans are as yet unclear. Hirai et al. [29] tried to clarify the neural dynamics in BM perception by comparing ERPs elicited by point-light displays of BM and scrambled motion. They report that both types of stimuli elicited peaks at around 200 and 240 ms which were larger in the BM condition than in the scrambled motion condition.

The aim of the present study was to further elucidate the nature, time course and location of neural processing involved in the perception of BM by using event-related potentials and, in addition, low resolution brain electromagnetic tomography (LORETA). Point-light displays of whole body motion in upright and inverted orientation served as stimuli in order to focus on the dynamic aspects of body motion and to reduce form cues from body shape. In contrast to some previous ERP-studies [47,64] and in accordance to the approach of Hirai et al. [29], in the current study form information has to be derived from the information of the motion trajectories. Furthermore, the comparison of upright and inverted biological motion aims to provide deeper insight into distinct processing stages associated with BM since inverted displays convey the same structural information as upright BM but the detection of an actor is substantially impaired as shown in several psychophysical experiments [43,53,59]. Control stimuli were displays of scrambled motion in which dots have the same motion vector as in the BM condition but with their initial position being randomized. We predicted an inversion effect for biological motion in the time window up to 200 ms after stimulus onset as usually found for faces or other stimulus categories for which the observer is an expert. Moreover, we expected an ERP source specific to upright BM in the STS-complex, concerned with the fine analysis of motion patterns which provide biologically relevant information and contribute to social perception.

2. Methods

2.1. Participants

15 healthy volunteers (8 females, 7 males; ages 20–35 years) participated in this study, which was undertaken with the understanding and written consent of each participant. The procedure was approved by the Ethics Committee of the Ruhr-University Bochum. All participants had normal or corrected-to-normal vision.

2.2. Stimuli

The visual stimuli used in this study were obtained from 20 men and 20 women walking on a treadmill, which served as models to acquire BM data. Data were recorded in 3D space using a motion capture system (Vicon; Oxford Metrics, Oxford, UK). A framework which allows using linear methods to transform BM data [58] was applied. As result of this procedure, an "average walker" was computed from our data set and animated as a point-light display. The dots representing the major joints of the body were located at the ankles, the knees, the hips, the wrists, the elbows, the shoulders, the center of the pelvis, on the sternum, and in the center of the head.

The point-light displays were presented in frontal view on a black screen either in upright orientation, inverted orientation (180° rotated in fronto-parallel plane) or as scrambled motion. Fig. 1 illustrates the three categories of visual stimuli. In the latter condition, the moving dots had the same local motion trajectories as in the upright BM displays, but their initial starting position was randomized destroying the spatial relation among the dots. The area in which the scrambling occurred was matched with respect to size to the other stimulus conditions.

2.3. Experimental setup

Participants were seated in a dimly lit sound attenuated cabin, with response buttons under their right hands. A computer screen was mounted at a distance of 90 cm in front of the participant's eyes. At this distance the stimuli subtended a visual angle of approximately 4.1° in height and 1.6° in width. All stimuli were presented for 800 ms at the center of the

Fig. 1. Categories of stimuli: BM in upright orientation (BM), BM in inverted orientation (IBM) and scrambled motion (SCR).

screen; successive trials were separated by intertrial intervals of 2000 ms in which a black screen was presented. Stimulus and motion onset occurred simultaneously. The experiment consisted of 60 trials per condition, resulting in 180 trials altogether which were presented in randomized order. Participants were asked to maintain central eye fixation during the trials and to respond as quickly and accurately as possible by pressing the right button to dot patterns representing BM (in upright and inverted orientation) and the left button to scrambled motion. Observers did not receive feedback on their responses. Before starting the experimental trials observers were shown some demonstration trials of all stimulus conditions in order to familiarize them with the display and the setup.

2.4. EEG-recording

EEG was recorded with Ag-AgCl electrodes mounted in an elastic cap from 30 scalp sites (F5, FZ, F6, T7, C5, C3, CZ, C4, C6, T8, TP7, CP5, CP3, CP4, CP6, TP8, P7, P5, P3, PZ, P4, P6, P8, PO7, O1, OZ, O2, PO8, A1, A2) according to the 10–20 system, referenced to an electrode on the tip of the nose. EOG was recorded from above and below the left eye as well as from the outer canthi of both eyes. Impedance was kept below 5 kΩ. A Neuroscan Synamps System with related software was used for recording. EEG was sampled with 200 Hz and stored on hard disk.

The EEG data were analyzed off-line using the Brain Vision Analyzer software package. Raw data were digitally filtered with a 0.1 Hz high-pass and a 40 Hz low-pass filter and segmented into epochs ranging from 200 ms before stimulus onset to 2000 ms after stimulus onset. After removing segments containing artifacts, ocular correction was carried out. Artifact detection was automatically performed (criterion $\pm 75\,\mu V$) and visually checked, afterwards. Epochs were baseline-corrected using the signal during the 200 ms that preceded the onset of the stimulus and averaged according to the three experimental conditions. Thereafter, grand averaged ERPs were calculated. Since the error rate was generally very low (less than 5%), all trials were included in the analysis.

2.5. Data analysis

Behavioral performance was analyzed by conducting a repeated measure one-way ANOVA to determine effects of stimulus condition on response times. ERP effects of experimental variables were determined by conducting repeated measures ANOVAs on ERP peak amplitude values (N170) or ERP mean amplitude values (N300). N170 amplitude was measured as peak amplitude in the 150–200 ms time range using an automated procedure. N300 amplitude was defined as mean activity in the time range between 230 and 360 ms after stimulus onset and calculated automatically. Peak latencies were analyzed for the N170 component and for the positive component preceding the N300 component in the time window 220–280 ms after stimulus onset. The ANOVAs
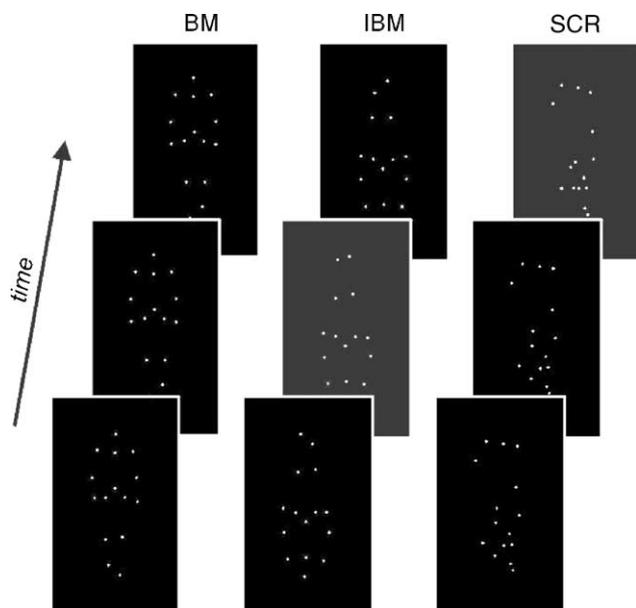
were conducted for the factors stimulus category (BM versus inverted BM versus scrambled motion), selected electrode locations (O1, O2 versus PO7, PO8 versus P7, P8 versus TP7, TP8) and recording side (right versus left). Electrode selection was based on the main areas of interest as suggested by previous functional imaging studies [11,24–27,51,62]. Greenhouse-Geisser adjustments to the degrees of freedom were performed when appropriate.

2.6. LORETA-analysis

LORETA [40,41] calculates the current density at each voxel in the gray matter and the hippocampus of a reference brain as a linear, weighted sum of the scalp electric potentials. LORETA chooses the smoothest of all possible current density configurations throughout the brain volume. This procedure only implicates that neighboring voxels should have a maximally similar activity, no other constraints are used. LORETA-images represent the electrical activity at each of the voxel as squared magnitude of the computed current density.

Amplitudes of the N170 component and mean activity between 230 and 360 ms (N300) were exported from the results as ASCII data for LORETA analysis. For each subject and each condition, one LORETA image was generated. These images were converted (http://www.ihb.spb. ru/ pet_lab/L2S/L2SMain.htm) for further analysis with SPM99 (http://www.fil.ion.ucl.ac.uk/spm/). In SPM99 a PET/SPECT design with a two sample t-test was performed. The following parameters were used for the analysis: global normalisation with proportional scaling and proportional scaling to a global mean = 50, absolute threshold masking with an analysis threshold set to 0 and global calculation of mean voxel value (within per image).

The level of significance was set to $P < 0.03$. Foci of significant differences were transformed into Talairach space [54] using the algorithm suggested by Brett (http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html) for anatomical labelling.

3. Results

3.1. Behavioral performance

Behavioral performance of correct identification of all three stimulus categories exceeded 95% across all conditions. RTs in the upright BM condition were significantly shorter than in the other conditions ($F(2,20) = 27.66$; $P < 0.001$). RTs in the inverted BM condition and in the scrambled motion condition did not differ significantly from each other as tested post hoc with Bonferroni adjusted measures (Fig. 2). RT analysis is based on 11 out of 15 subjects, because of missing data in four subjects.
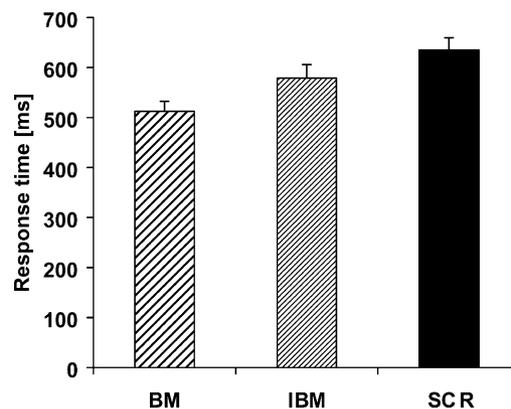

Fig. 2. Reaction times for correct responses in the three experimental conditions. Error bars indicate S.E.M.

3.2. ERP effects

Fig. 3A shows grand averaged ERPs in response to the three experimental conditions. In the latency window up to 400 ms after stimulus onset, two distinct components emerged in all experimental conditions. The first negative component peaks on average at a latency of 183 ms. In accordance with previous studies [46–48,64], this component was termed N170.

The later negative component is located in the time window between 230 and 360 ms after stimulus onset. We refer to this second negative component as N300. Fig. 3B shows the difference waveforms of BM minus scrambled motion as well as inverted BM minus scrambled motion to illustrate the differential amplitudes of BM and inverted BM compared to scrambled motion. These difference waveforms reach their largest amplitudes earlier than the N170 peak and the N300 peak, respectively.

3.3. N170 amplitude and latency

N170 amplitude was assessed individually as peak amplitude within the 150–200 ms latency window. ANOVA yielded significant main effects of stimulus condition ($F(2,28) = 9.97$; $P = 0.001$) and electrode location ($F(2,28) = 4.42$; $P = 0.044$) as well as a significant interaction ($F(2,28) = 3.31$; $P = 0.018$). Differences between stimulus conditions were tested post-hoc with Bonferroni adjusted measures yielding a significant difference between the BM and scrambled motion conditions ($P = 0.003$) and between the BM and inverted BM conditions ($P = 0.029$). The main effect of electrode locations is due to generally smaller peak amplitudes at parieto-temporal electrodes in comparison to the other sites. The interaction between stimulus condition and electrode location reflects larger peak amplitude differences between upright BM and the other conditions for posterior electrodes than for anterior electrodes.

Analysis of peak latencies for the N170 component (Table 1) did not yield significant differences between the three conditions ($F(2,28) = 0.32$; $P = 0.969$).
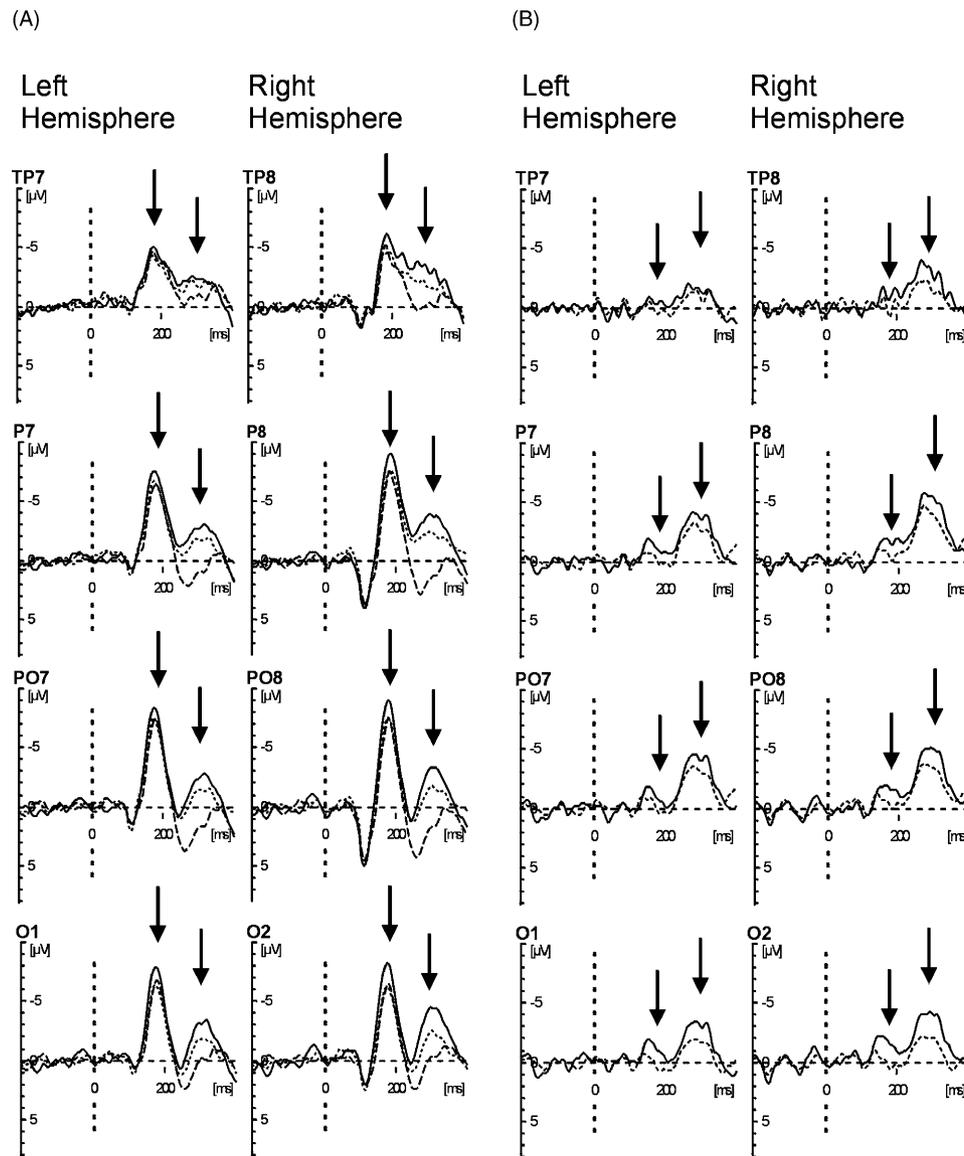
Fig. 3. (A) Grand-averaged ERPs recorded at lateral posterior electrodes (left hemisphere: O1, PO7, P7 and TP7; right hemisphere: O2, PO8, P8 and TP8) in response to BM (solid lines), inverted BM (dotted lines) and scrambled motion (dashed lines). Arrows indicate mean peak latency of the N170 and the N300 component. (B) Difference waveforms obtained by subtracting ERP's to scrambled motion from ERPs to BM (solid line) and by subtracting ERPs to scrambled motion from ERPs to inverted BM (dotted line).

3.4. N300 amplitude and latency

The later negative component (N300) was quantified as mean amplitude within the 230–360 ms latency window. The N300 component had a larger amplitude for upright and inverted walkers as compared to scrambled walkers ($F(2,28) =$

Table 1
Peak latencies for the N170 component and the positive peak preceding the N300 component (means and S.E.M.s in ms)

	BM	IBM	SCM
Mean N170 peak latency	184.3 (2.7)	181.6 (3.1)	182.9 (3.1)
Mean positive peak latency pre N300	246.1 (3.4)	247.7 (4.9)	257.9 (3.0)

14.90; $P < 0.001$). A post-hoc test with Bonferroni adjusted measures revealed a significant difference between the BM and scrambled motion conditions ($P = 0.001$) and between the inverted BM and scrambled motion conditions ($P = 0.003$). In addition, there was a significant interaction between stimulus condition and side of recording ($F(2,28) = 4.21$; $P < 0.027$). This interaction is due to larger differences in mean amplitudes between BM and scrambled motion at right temporo-parietal and parietal electrodes. This indicates a pronounced right-hemispheric advantage in visual processing of BM, particularly at later processing stages.

In addition to amplitude differences, significant differences in the latencies (Table 1) of the positive peak preceding the N300 component emerged ($F(2,28) = 4.70$; $P < 0.024$).

Table 2
Results of LORETA analysis for N170 and N300 contrasts showing Talairach space coordinates, probable Brodman areas (BA) in the range of 3 mm and levels of significance

	BA	x	y	z	P
N170 BM vs. SCR					
Lingual gyrus	10	−24	79	−7	0.016
Posterior cingulate gyrus	23, 30	−3	−44	22	0.024
N170 BM vs. IBM					
Subcallosal gyrus, precuneus		−16	−64	23	0.014
Posterior cingulate gyrus	30	−3	−51	16	0.014
Middle occipital gyrus	19	46	−78	11	0.024
N300 BM vs. SCR					
Fusiform gyrus, Cerebellum	6	32	−45	−15	0.001
Subcallosal gyrus, Anterior cingulate gyrus	6	−3	9	−11	0.009
Medial frontal gyrus, rectal gyrus	25, 11	−10	16	−19	0.015
Superior temporal gyrus		52	−37	16	0.011
N300 IBM vs. SCR					
Inferior frontal gyrus, middle frontal gyrus	47, 11	25	29	−12	0.004
Medial frontal gyrus, rectal gyrus	25, 11	−3	16	−18	0.013
Anterior cingulate gyrus	25	−3	16	−6	0.013
Superior temporal gyrus		52	−37	9	0.027

The peak was delayed in the scrambled motion condition, whereas the latencies for BM and inverted BM did not differ significantly. Moreover, there was also a significant effect of peak amplitudes for this positive peak ($F(2,28) = 18.76$; $P < 0.001$) with a pronounced positivity for scrambled motion. Amplitude differences between BM and inverted BM were not significant.

3.5. Source analysis

Source analysis was performed for two reasons: to consider the data from all electrodes during the two time windows of interest and to get an approximation of the ERP component sources. Results of the LORETA analysis are listed in Table 2. It has to be pointed out that the spatial resolution of this analysis is not as high as resolution from fMRI experiments. Source analysis was performed separately for the N170 and N300 component.

For the N170 component, the contrast between upright BM and scrambled motion and the contrast between upright BM and inverted BM were calculated, since the ERP amplitudes between those conditions differed significantly from each other. The contrast between upright BM and scrambled motion revealed sources in the posterior cingulate gyrus and in the left lingual gyrus. Results for the upright BM versus inverted BM contrast yielded three distinct sources: one in the posterior cingulate gyrus, one in the area of the subcallosal gyrus/precuneus and another in the right occipital gyrus. Fig. 4 illustrates schematically the location of the N170 sources for both contrasts.

For the N300 component, the contrast between upright BM and scrambled motion and the contrast between inverted BM and scrambled motion were calculated. The largest contrast between upright BM and scrambled motion emerged for the right fusiform gyrus. In addition, a source within the right

superior temporal gyrus was estimated. Additional sources were observed in the orbitofrontal cortex (subcallosal gyrus and anterior cingulate gyrus; rectal gyrus and medial frontal gyrus). Results for the inverted BM versus scrambled motion contrast yielded four distinct sources. One source was located in the right superior temporal gyrus, the other sources were located in orbitofrontal brain areas (rectal gyrus and medial
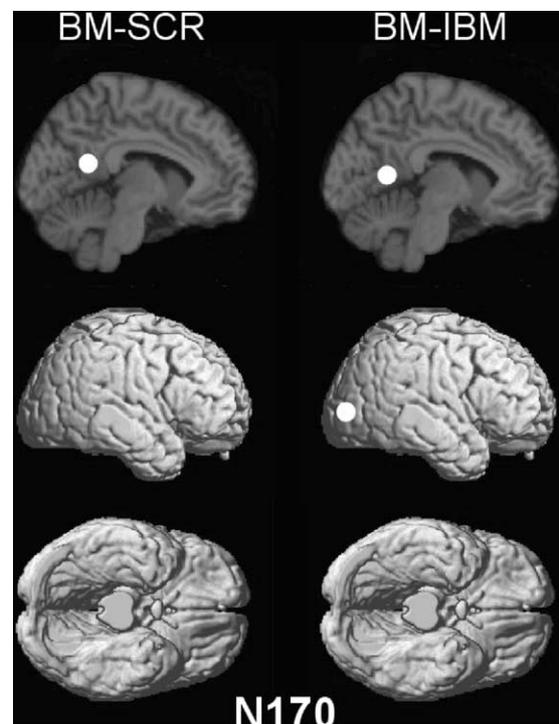

Fig. 4. LORETA-analysis: group comparison of absolute current density values between the upright BM condition and the scrambled motion condition (BM-SCR) and between the upright BM condition and the inverted BM condition (BM-IBM) for the N170 peak. Three percent *P*-value threshold.
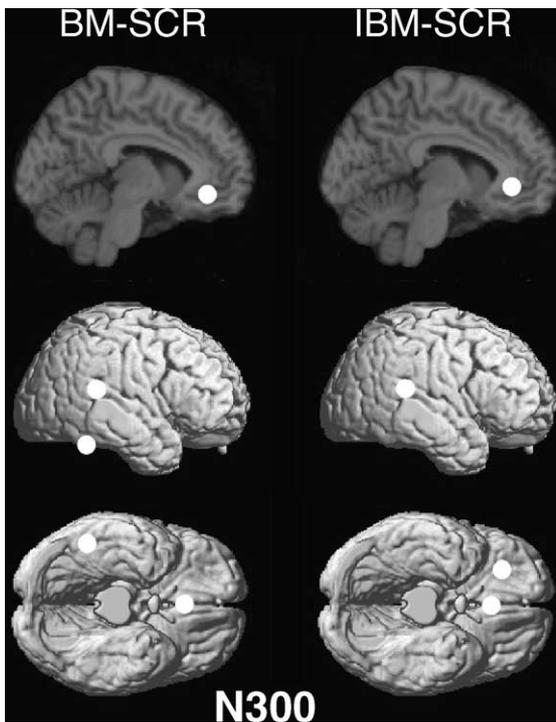
Fig. 5. LORETA-analysis: group comparison of absolute current density values between the upright BM condition and the scrambled motion condition (BM-SCR) and between the inverted BM condition and the scrambled motion condition (IBM-SCR) in the time range 230–360 ms after stimulus onset (N300). Three percent *P*-value threshold.

frontal gyrus; and anterior cingulate gyrus) and in the Inferior frontal gyrus. Locations of the N300 sources for both contrasts are illustrated in Fig. 5.

LORETA analysis yielded evidence for sources specific to BM within the right fusiform gyrus and the right superior temporal gyrus for the second component, whereas sources specific for BM in the early component were located in areas associated with attentional aspects of visual processing (posterior cingulate cortex). Additional sources generating the second component were located in orbitofrontal brain areas (anterior cingulate gyrus, medial frontal gyrus).

4. Discussion

The current results present clear evidence for the involvement of two distinct processing stages in the visual analysis of BM: an early negative component (N170) peaking at 180 ms after stimulus onset and a later negative component (N300) in the time window between 230 and 360 ms after stimulus onset. The N170 component is modulated differently by upright BM in comparison to inverted BM and scrambled motion. The difference between upright and inverted BM reflects an inversion effect in BM perception. The amplitude of the N300 component did not differ significantly between inverted and upright BM stimuli, but was less pronounced for scrambled motion, indicating similar processing for upright and in-

verted BM conditions in later processing stages. Whereas the sources generating the N170 component are mainly located in posterior areas near the midline, the later N300 component is generated by sources in the superior temporal gyrus and the fusiform gyrus in the right hemisphere.

In our experimental setup, onset of stimulus presentation and onset of motion occur in parallel. Therefore, the ERPs are evoked in response to stimulus onset as well as in response to motion onset and consequently both processes may contribute to the neural responses recorded in the ERPs. The neural basis of motion perception has been studied psychophysiologically using motion-onset VEPs [3,4,31]. Visual motion onset evokes VEP at two major sites, the occipital/occipital-temporal sites and the vertex. At occipital sites, visual motion onset per se evokes ERP components that are dominated by a minor positivity (P1) around 100–130 ms and a pronounced negativity (N2) around 150–200 ms. The negative component represents motion mechanisms as shown by its susceptibility to motion adaptation, while the positive component is more likely associated with form-processing mechanisms. In visual perception of faces or man-made objects like houses, stimulus onset elicits a negative component around a latency of 170 ms [21]. Taking these factors into consideration, the first negative ERP component (N170) obtained in the present experiment may reflect the contribution of both processes. Since both processes are inherent features of BM, the relative contribution of each process to ERPs is difficult to estimate.

BM processing occurs with very short latencies. Visual processing needed to perform the highly demanding task of discrimination between upright BM patterns on the one hand and scrambled motion and inverted BM on the other hand can be achieved within a time period of 180 ms. In static images of natural scenes processing required to decide whether the scene contains an animal can even be performed within 150 ms, as measured by ERP modulation [57]. Processing speed cannot be improved by extensive training [22], and seems to be limited by the underlying neural mechanism. The present study is concerned with motion rather than with static stimuli representing animate versus inanimate stimuli. Given a monitor frame rate of 60 Hz, the visual system needs at least $2/60 \sim 30$ ms to integrate two frames, which is a prerequisite for generation of a coherent percept in stimuli of BM. Because of the short latencies obtained in the present study, it seems likely that the processing of BM must be based on highly automated feed-forward mechanisms.

Nevertheless, we cannot rule out completely that static cues have played a minor role to perform the discrimination task. The three classes of stimuli can be discriminated on the basis of static versions of the displays since the spatial arrangement is slightly different for upright, inverted and scrambled stimuli. Such a discrimination of static displays would be based purely on geometrical cues and is clearly different from the discrimination based on the percept of a walking human figure which is only evoked by animated displays of biological motion. Moreover, we found sources of the ERP signal in the second component in brain areas which

were reported to be selectively involved in biological motion perception in several imaging studies [11,24–27,51,62] but were not reported to be activated by static displays consisting of a few dots which match joint positions.

The ERP effects are in accordance with psychophysical findings showing a strong inversion effect in perception of BM [9,18,19,37,43,53,59] as well as in face perception, e.g. [55,59,63]. As expected, the ERP effects are mirrored by behavioral data, showing shorter response times for the detection of upright walkers in comparison to inverted walkers and scrambled motion. Taking into account the longer response times for inverted BM and scrambled motion, similar modulations in the early ERP-component in response to inverted BM and scrambled motion can be expected. The pronounced effect for upright BM, therefore, probably reflects the pop-out effect of BM in familiar orientation that leads to shorter reaction times for the detection of this type of visual information. This pop-out effect might be associated with the global recognition of a human person depicted as a point-light display as previously described [10].

A pop-out phenomenon resulting from perceptual experience requires the involvement of high-level areas, since complex visual stimuli such as point-light displays of BM must be integrated in neural populations having large receptive fields. Due to the short latency of the N170 component, this can only be achieved by a feedforward mechanism. This interpretation is in accordance with the reverse hierarchy theory [30] suggesting that "vision at a glance" matches a high-level, generalized, categorical scene interpretation by neural processing along the feedforward hierarchy of areas leading to increasingly complex representations. High-level spread attention associated with "vision at a glance" subserves the initial, crude global percept of the gist of a scene. The pop-out effect is only one aspect of this crude initial assessment. In contrast, for later "vision with scrutiny", reverse hierarchy routines focus attention to specific units incorporating detailed information available there into conscious perception. Therefore, fine discrimination depends on re-entry to low-level specific receptive fields to bind features.

In the N170 component, sources generating both the contrast between BM and inverted BM and the contrast between BM and scrambled motion are located in the posterior cingulate cortex. This area seems to have different attention-related functions. It has been suggested that the cingulate region may establish a neural interface between attention and motivation [52]. In addition, activity in the posterior cingulate cortex was correlated with the speed of detecting a visual target when it was preceded by a predictive cue [36]. We assume that the posterior cingulate cortex might reflect high-level spread attention subserving neural processing leading to the global percept. The finding of sources generating the contrast between BM and the other conditions in attention-related areas also fits well with recent behavioral data suggesting that attention is required for the visual analysis of point-light displays [14,56]. Cavanagh et al. [14] showed that discrimination of specific features of point-light displays of BM seems to be a

serial process since reaction times increased with the number of items. The reaction time increase was attributed to increasing attentional demands of the task. Results in a dual task paradigm to explore the role of attention in the processing of BM [56] suggested that, in some cases, perception of BM can be automatic. But if strategies operating in a global, top-down fashion are required, attentional demands play a vital role.

The N300 component might reflect processes associated with "vision with scrutiny" according to the reverse hierarchical theory [30] and is responsible for the fine analysis of BM patterns which is necessary to retrieve visual information of social and psychological relevance. This view is in agreement with source localization relating to the contrast between BM and scrambled motion in the second component. These sources are located in the right superior temporal gyrus and the right fusiform gyrus. For the contrast between inverted BM and scrambled motion, there was only one source located in the superior temporal gyrus. In several brain imaging studies, the STS complex as well as the fusiform face area were shown to be involved in perception of BM. [11,24–27,51,62]. The present results from ERPs and source analysis are in accordance with the findings from these brain imaging studies.

In addition to STG and FFG, we also found sources within the anterior cingulate cortex and the orbital prefrontal cortex for the second component. Several studies have shown that the anterior cingulate cortex is involved in response selection [12], monitoring of performance and conflict evaluation [13] and regulation of attention [45]. The anterior cingulate cortex source may be thus more related to cognitive processes associated with the task requirements (to make a decision between two response alternatives) than to perceptual analysis of various stimulus classes. A number of studies [1] provided evidence that the orbitofrontal cortex plays a crucial role in social cognition and has strong interconnections to the STS-complex as well as to the fusiform gyrus which are engaged in social perception. In the present study, sources in the orbitofrontal cortex might also be related to these neural mechanisms.

The ERP correlate of the inversion effect in BM perception are opposite to inversion effects obtained in face recognition paradigms, in which inverted faces elicit a higher negativity in the N170 component than upright faces [49]. When comparing upright faces with control stimuli (consisting of houses, etc.), faces elicit a more pronounced N170 effect [20,21]. In the present procedure, we used scrambled motion as control stimuli. There were no differences in the early component between scrambled and inverted BM, indicating that these two stimulus types are processed similarly during initial processing stages. In addition, the mean peak latency of the first component was 180 ms. This latency is longer than the peak latencies usually reported for face perception [8]. This difference may be a consequence of an extended integration time for the detection of form conveyed by motion which is not required for static image perception.

The finding of hemispheric asymmetries in STG and FFG in later processing of BM representing human gait patterns is also consistent with evidence from fMRI studies [24,25], which reported pronounced activity in the right hemisphere associated with the perception of BM. This finding may be related to information useful for the recognition of individual human features. These displays contain such sources of information and provide important cues for social interaction.

Our results are in accordance with the ERP – findings from Hirai et al. [29] who reported peaks at around 200 and 240 ms after stimulus onset which were larger in response to BM than to scrambled motion. In contrast to our study, they used only two experimental conditions (BM versus scrambled motion) and did not include source analysis. Their displays were presented in profile view and were generated by a computer algorithm. There are, however, differences concerning the latency of the first component which was on average shorter in our paradigm than in the study by Hirai et al. [29]. Our study extends this work by demonstrating inversion effects in early visual processing reflecting an pop-out effect with upright walkers, but not in later processing stages associated with the fine analysis of BM. In addition, our findings offer some evidence for the brain areas associated with the two components.

A recently published paper [42] analyzed gamma MEG activity in response to BM generated from a computer algorithm. Recognizable upright and non-recognizable inverted walkers evoked enhancements in oscillatory gamma brain activity (25–30 Hz) over the left occipital cortices as early as 100 ms from stimulus onset. Upright BM elicited further gamma response over the parietal (130 ms) and right temporal (170 ms) lobes. Whereas the temporal order and approximate localization of brain areas showing synchronized firing pattern in the gamma band is in accordance with our findings, the absolute timing of the occurrence of gamma activity is different from the timing of the ERP-components found in our study. This difference might be related to differences in recording techniques (MEG versus EEG) and analysis methods (frequency analysis versus ERP).

Taken together, our findings suggest two distinct components of BM processing, recruiting different neuronal populations. The first processing stage (N170) reflects the generation of a global percept of the visual scene leading to a pop-out effect of upright biological motion. The sensitivity to upright biological motion is the result of the familiarity of BM in normal orientation. At the level of the second processing stage (N300), brain areas as STG and FFG, which are known from fMRI studies to be involved in the fine and detailed perceptual analysis of BM, play an important role. The evidence for fast, efficient processing underlines the importance of perception of BM, and provides further evidence for a specific neural network involved in processing biologically relevant motion signals. The right-hemispheric dominance associated with BM perception shows clear parallels to asymmetries in face perception and probably reflects the social relevance of animate motion perception. Furthermore, STG and FFG are primarily involved in the second ERP component (N300) and show clear hemispheric asymmetries in the perceptual analysis of BM only during later processing stages.

Acknowledgments

References

[1] Adolphs R. The neurobiology of social cognition. Curr Opin Neurobiol 2001;11(2):231–9.

[2] Allison T, Puce A, McCarthy G. Social perception from visual cues: role of the STS region. Trends Cogn Sci 2000;4(7):267–78.

[3] Bach M, Ullrich D. Motion adaptation governs the shape of motion-evoked cortical potentials. Vision Res 1994;34(12):1541–7.

[4] Bach M, Ullrich D. Contrast dependency of motion-onset and pattern-reversal VEPs: interaction of stimulus type, recording site and response component. Vision Res 1997;37(13):1845–9.

[5] Barclay CD, Cutting JE, Kozlowski LT. Temporal and spatial factors in gait perception that influence gender recognition. Percept Psychophys 1978;23:145–52.

[6] Beardsworth T, Buckner T. The ability to recognize oneself from a video recording of one's movements without seeing one's body. Bull Psychonom Soc 1981;18:19–22.

[7] Beauchamp MS, Lee KE, Haxby JV, Martin A. FMRI responses to video and point-light displays of moving humans and manipulable objects. J Cogn Neurosci 2003;15(7):991–1001.

[8] Bentin S, Allison T, Puce A, Perez E, McCarthy G. Electrophysiological studies of face perception in humans. J Cogn Neurosci 1996;8(6):551–65.

[9] Bertenthal B, Proffitt DR, Kramer SJ. Perception of biological motion by infants: implementation of various processing constraints. J Exp Psychol: Hum Percept Perform 1987;13:577–85.

[10] Bertenthal BI, Pinto J. Global processing of biological motions. Psychol Sci 1994;5(4):221–5.

[11] Bonda E, Petrides M, Ostry D, Evans A. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. J Neurosci 1996;16(11):3737–44.

[12] Bunge SA, Hazeltine E, Scanlon MD, Rosen AC, Gabrieli JD. Dissociable contributions of prefrontal and parietal cortices to response selection. Neuroimage 2002;17(3):1562–71.

[13] Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 1998;280(5364):747–9.

[14] Cavanagh P, Labianca AT, Thornton IM. Attention-based visual routines: sprites. Cognition 2001;80(1–2):47–60.

[15] Cowey A, Vaina LM. Blindness to form from motion despite intact static form perception and motion detection. Neuropsychologia 2000;38(5):566–78.

[16] Cutting JE. Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. Perception 1978;7(4):393–405.

[17] Cutting JE, Kozlowski LT. Recognizing friends by their walk: Gait perception without familiarity cues. Bull Psychonom Soc 1977;9(5):353–6.

[18] Dittrich WH. Action categories and the perception of biological motion. Perception 1993;22(1):15–22.

[19] Dittrich WH, Troscianko T, Lea SEG, Morgan D. Perception of emotion from dynamic point-light displays represented in dance. Perception 1996;25:727–38.

[20] Eimer M. Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. Brain Res Cogn Brain Res 2000;10(1–2):145–58.

[21] Eimer M. Event-related brain potentials distinguish processing stages involved in face perception and recognition. Clin Neurophysiol 2000;111(4):694–705.

[22] Fabre-Thorpe M, Delorme A, Marlot C, Thorpe S. A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. J Cogn Neurosci 2001;13(2):171–80.

[23] Giese MA, Poggio T. Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci 2003;4(3):179–92.

[24] Grezes J, Fonlupt P, Bertenthal B, Delon-Martin C, Segebarth C, Decety J. Does perception of biological motion rely on specific brain regions? Neuroimage 2001;13(5):775–85.

[25] Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, et al. Brain areas involved in perception of biological motion. J Cogn Neurosci 2000;12(5):711–20.

[26] Grossman ED, Blake R. Brain activity evoked by inverted and imagined biological motion. Vision Res 2001;41(10–11):1475–82.

[27] Grossman ED, Blake R. Brain areas active during visual perception of biological motion. Neuron 2002;35(6):1167–75.

[28] Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends Cogn Sci 2000;4(6):223–33.

[29] Hirai M, Fukushima H, Hiraki K. An event-related potentials study of biological motion perception in humans. Neurosci Lett 2003;344(1):41–4.

[30] Hochstein S, Ahissar M. View from the top: hierarchies and reverse hierarchies in the visual system. Neuron 2002;36(5):791–804.

[31] Hoffmann MB, Unsold AS, Bach M. Directional tuning of human motion adaptation as reflected by the motion VEP. Vision Res 2001;41(17):2187–94.

[32] Johansson G. Visual perception of biological motion and a model for its analysis. Percept Psychophys 1973;14(2):201–11.

[33] Kozlowski LT, Cutting JE. Recognizing the sex of a walker from a dynamic point-light display. Percept Psychophys 1977;21(6):575–80.

[34] Mather GML. Gender judgements in biological motion perception. Proc R Soc London 1994:273–9.

[35] McLeod P, Dittrich W, Driver J, Perrett D, Ziehl J. Preserved and impaired detection of structure from motion by a "motion-blind" patient. Visual Cogn 1996;3(4):363–91.

[36] Mesulam MM, Nobre AC, Kim YH, Parrish TB, Gitelman DR. Heterogeneity of cingulate contributions to spatial attention. Neuroimage 2001;13(6 Part 1):1065–72.

[37] Mitkin AA, Pavlova MA. Changing a natural orientation: recognition of biological motion pattern by children and adults. Psychologische Beitraege 1990;32(1–2):28–35.

[38] Oram MW, Perrett DI. Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. J Cogn Neurosci 1994;6(2):99–116.

[39] Oram MW, Perrett DI. Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. J Neurophysiol 1996;76(1):109–29.

[40] Pascual-Marqui RD, Lehmann D, Koenig T, Kochi K, Merlo MC, Hell D, et al. Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia. Psychiatr Res 1999;90(3):169–79.

[41] Pascual-Marqui RD, Michel CM, Lehmann D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int J Psychophysiol 1994;18(1):49–65.

[42] Pavlova M, Lutzenberger W, Sokolov A, Birbaumer N. Dissociable cortical processing of recognizable and non-recognizable biological movement: analysing gamma MEG activity. Cereb Cortex 2004;14(2):181–8.

[43] Pavlova M, Sokolov A. Orientation specificity in biological motion perception. Percept Psychophys 2000;62(5):889–99.

[44] Pelphrey KA, Mitchell TV, McKeown MJ, Goldstein J, Allison T, McCarthy G. Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. J Neurosci 2003;23(17):6819–25.

[45] Posner MI, DiGirolamo GJ. Executive attention: conflict, target detection, and cognitive control. In: Parasuraman R, editor. The attentive brain. Cambridge, MA: MIT Press; 1998. p. 401–23.

[46] Puce A, Perrett D. Electrophysiology and brain imaging of biological motion. Philos Trans R Soc Lond B Biol Sci 2003;358(1431):435–45.

[47] Puce A, Smith A, Allison T. ERPs evoked by viewing facial movements. Cogn Neuropsychol 2000;17(1–3):221–39.

[48] Puce A, Syngeniotis A, Thompson JC, Abbott DF, Wheaton KJ, Castiello U. The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. Neuroimage 2003;19(3):861–9.

[49] Rossion B, Gauthier I. How does the brain process upright and inverted faces? Behav Cogn Neurosci Rev 2002;1(1):62–74.

[50] Schenk T, Zihl J. Visual motion perception after brain damage. II. Deficits in form-from-motion perception. Neuropsychologia 1997;35(9):1299–310.

[51] Servos P, Osu R, Santi A, Kawato M. The neural substrates of biological motion perception: an fMRI Study. Cereb Cortex 2002;12(7):772–82.

[52] Small DM, Gitelman DR, Gregory MD, Nobre AC, Parrish TB, Mesulam MM. The posterior cingulate and medial prefrontal cortex mediate the anticipatory allocation of spatial attention. Neuroimage 2003;18(3):633–41.

[53] Sumi S. Upside-down presentation of the Johansson moving light-spot pattern. Perception 1984;13(3):283–6.

[54] Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. Thieme Medical Publishers; 1988.

[55] Thompson P. Margaret Thatcher – a new illusion. Perception 1980;9:483–4.

[56] Thornton IM, Rensink RA, Shiffrar M. Active versus passive processing of biological motion. Perception 2002;31(7):837–53.

[57] Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. Nature 1996;381(6582):520–2.

[58] Troje NF. Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. J Vision 2002;2(5):371–87.

[59] Troje NF. Reference frames for orientation anisotropies in face recognition and biological-motion perception. Perception 2003;32(2):201–10.

[60] Vaina LM. Functional segregation of color and motion processing in the human visual cortex: clinical evidence. Cereb Cortex 1994;4(5):555–72.

[61] Vaina LM, Lemay M, Bienfang DC, Choi AY, Nakayama K. Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: a case study. Vis Neurosci 1990;5(4):353–69.

[62] Vaina LM, Solomon J, Chowdhury S, Sinha P, Belliveau JW. Functional neuroanatomy of biological motion perception in humans. Proc Natl Acad Sci USA 2001;11:11.

[63] Valentine T. Upside-down faces: a review of the effect of inversion upon face recognition. Br J Psychol 1988;79(Part 4):471–91.

[64] Wheaton KJ, Pipingas A, Silberstein RB, Puce A. Human neural responses elicited to observing the actions of others. Vis Neurosci 2001;18(3):401–6.