NeuroImage 84 (2014) 854–867

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Electrophysiological markers of biological motion and human form recognition

Nicole C. White [a,1], Jonathan M. Fawcett [a,2], Aaron J. Newman [a,b,c,d,*]

[a] Department of Psychology and Neuroscience, Life Sciences Centre, Dalhousie University, Halifax, NS B3H 4R2, Canada
[b] Department of Psychiatry, Dalhousie University, Halifax, NS, Canada
[c] Department of Surgery, Dalhousie University, Halifax, NS, Canada
[d] Department of Pediatrics, Division of Neurology, Dalhousie University, Halifax, NS, Canada

ARTICLE INFO

Article history:
Accepted 13 September 2013
Available online 21 September 2013

ABSTRACT

Current models suggest that human form and motion information are initially processed through separate pathways, then integrated in action perception. Testing such a sequential model requires techniques with high temporal resolution. Prior work demonstrated sensitivity of a posterior temporal event-related potential (ERP) effect — the N2 — to biological motion, but did not test whether the N2 indexes biological motion perception specifically, or human form/action perception more generally. We recorded ERPs while participants viewed stimuli across 3 blocks: (1) static (non-moving) point-light displays of humans performing actions; (2) static stick figures with clear forms; and (3) point-light biological motion. A similar sequence of ERP components was elicited by human forms in all blocks (stationary and moving), and reliably discriminated between human and scrambled forms. The N2 showed similar scalp distribution and sensitivity to stimulus manipulations for both stick figures and biological motion, suggesting that it indexes integration of form and motion information, rather than biological motion perception exclusively — and that form and motion information are therefore integrated by approximately 200 ms. We identified a component subsequent to the N2, which we label the medial parietal positivity/ventral–anterior negativity (MPP/VAN), that was also sensitive to both human form and motion information. We propose that the MPP/VAN reflects higher-order human action recognition that occurs subsequent to the integration of form and motion information reflected by the N2.

Introduction

The human brain is highly sensitive to the movements of biological (i.e., living) entities. In particular, human actions can be recognized accurately from stimuli containing little to no detailed information about the human form. Johansson (1973) used point-light stimuli, consisting of approximately a dozen points of light (point-lights) located on the major joints of the human body, to demonstrate that explicit form information was unnecessary for recognizing human movements. More recent work has demonstrated that biological motion engages specialized brain mechanisms distinct from those used for motion perception more generally. Selective brain damage can impair low-level motion information but spare biological motion perception (Vaina et al., 1990), and manipulating low-level cues to motion information (luminance and contrast) differentially affects biological and low-level motion perception (Garcia and Grossman, 2008). Further, manipulating local motion information (individual point-light trajectories) does not impair biological motion recognition, but reducing the amount of form information present in the stimulus does (Beintema and Lappe, 2002; Beintema et al., 2006).

Thus while explicit form information is not required for biological motion perception, the data suggest that knowledge of shape/form information is involved in biological motion perception. While lower-level motion processing is dependent on area V5 in the middle temporal sulcus, lesion and neuroimaging studies have implicated the posterior superior temporal sulcus (STSp) as critical to the perception of biological motion (Beauchamp et al., 2002; Bonda et al., 1996; Grossman et al., 2000; Saygin, 2007; Servos et al., 2002; Vaina and Gross, 2004; Vaina et al., 2001), as well as the posterior parietal cortex, inferior temporal regions on the fusiform and lingual gyri, and premotor areas. It has been suggested that STSp is an area where motion and form information, processed somewhat separately by the dorsal and ventral visual streams, respectively, are integrated (see Blake and Shiffrar, 2007 for a review). The functional role of STSp is sometimes contrasted with nearby lateral temporal–occipital regions, the extrastriate body area (EBA) and fusiform body area (FBA), which show selectivity for human body forms over other static images (Downing et al., 2001; Peelen et al., 2006). However, while STSp may show stronger selectivity for moving

* Corresponding author at: Department of Psychology and Neuroscience, Life Sciences Centre, Dalhousie University, Halifax, NS B3H 4R2, Canada.
E-mail address: Aaron.Newman@Dal.ca (A.J. Newman).
URL: http://neuroimaging.psychology.dal.ca (A.J. Newman).
[1] Now at the University of Toronto.
[2] Now at the MRC Cognition and Brain Sciences Unit, Cambridge.

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.09.026

than static forms, and EBA the reverse pattern, both STSp and EBA show activity for both static and moving forms (Grossman and Blake, 2002; Peelen et al., 2006). For example, when participants are forced to rely more on form information for biological motion recognition, stronger activation in the EBA is observed (Michels et al., 2005), and much of the area defined as EBA overlaps with parts of the motion-sensitive V5 complex (Ferri et al., 2013). Jastorff and Orban (2009) suggested that EBA and FBA, rather than being selectively sensitive to shape information, integrate form and motion information while STSp provides more in-depth analysis of human actions. Collectively, these data suggest that form and motion information are used in combination during normal perception of biological motion.

A computational model of biological motion recognition, based on neurophysiological data, proposes that biological motion perception occurs through an integration of "snapshots" of static forms (Lange and Lappe, 2006). Similarly, a recent model based on a review of the neuropsychological and neuroimaging literature proposed that form and motion information are integrated in early extrastriate regions before biological motion is recognized, but that the relative importance of each may be dynamic depending on the information available and, in neuropsychological cases, the location and extent of brain lesions (Matheson and Mcmullen, 2010). This model proposes that form and motion information are first processed independently, then integrated. However, this hypothesis cannot be tested using neuropsychological or fMRI data, as those techniques lack the required temporal resolution. Event-related potentials (ERPs) have the ability to resolve the time course of neural activation. However, while biological motion processing has been studied with ERPs, none of these studies have compared biological motion to static form recognition. The present study examined the time course of human form and motion identification, comparing ERPs elicited by both static and moving human figures to assess whether biological motion recognition elicits a unique electrophysiological signature relative to static human forms.

ERPs associated with biological motion perception

A handful of previous studies have investigated biological motion perception using ERPs, and two components have been suggested to show sensitivity to biological motion, both maximal over posterior occipito-temporal scalp regions: a negativity peaking around 170–200 ms post-stimulus onset, and another negative peak between 200 and 300 ms. The relative timing of these two components varies across studies, as do the labels assigned to them. The first negativity is variously labeled the N1 Krakowski et al., 2011, N170 (Jokisch et al., 2005), N200 (Hirai et al., 2003, 2005), or N210 (Hirai and Hiraki, 2006). The second negativity has likewise been labeled the N2 (Krakowski et al., 2011), N240 (Hirai et al., 2003), N280 (Hirai and Hiraki, 2006), N300 (Jokisch et al., 2005), and N330 (Hirai et al., 2005). A summary of these findings and their eliciting conditions is given in Table 1. Given the consistency with which these two waveform peaks occur across studies, in terms of the scalp topography, polarity, eliciting conditions, and approximate timing, it seems reasonable to interpret these effects as reflecting two subsequent stages of neurocognitive processing across studies. Here we will use the convention of labeling these differences in order of their occurrence rather than their absolute timing, referring to these negative peaks as the N1 and N2. However, two things should be kept in mind regarding this naming convention. Firstly, in the literature there are multiple ERP components that have been labeled "N1" and "N2" in different studies and experimental contexts, with different scalp distributions, latencies, and eliciting conditions. Here we remain neutral with respect to how the N1 and N2 effects elicited by biological motion stimuli relate to these other components, though we return to this in the discussion. Secondly, increased negativities observed in previous studies of biological motion have not necessarily corresponded directly to the peak negative values observed in the ERP waveforms. That is to say, the enhanced negativities sometimes span both a negative and positive waveform peak (e.g., N1 and P2). Thus an "N1" difference between conditions should not be equated with an effect limited in timing to the N1 waveform peak. In the present paper we will use the unqualified labels "N1" and "N2" to refer to between-condition differences, and use terms such as "N1 waveform peak" when referring to the peaks and troughs observable in the waveforms from individual conditions.

All of the reported studies have contrasted upright point-light biological motion with scrambled point-lights (preserving the motion vectors but randomizing the starting position of each point, to disrupt the global percept of a human form); one study also included an inverted biological motion condition (Jokisch et al., 2005). The findings across studies are fairly consistent, with upright biological point-light motion eliciting stronger negativities than scrambled motion in both the N1 and N2 time windows (with one exception: Hirai and Hiraki, 2006 did not find significant differences between conditions for the first negative peak). Jokisch et al. (2005) found a larger N1 for upright than inverted biological motion, and no N1 differences between inverted and scrambled motion. For the N2, however, they found that inverted biological motion elicited a similar response as did upright biological motion, with both being more negative than for scrambled motion.

Thus both the N1 and N2 differentiate canonical (upright) biological motion from scrambled motion, however the N1 effect is selective for upright biological motion, while the N2 additionally responds to inverted biological motion. This finding may indicate that an initial stage of processing point-light stimuli is uniquely sensitive to upright forms, while a later stage is able to also recognize inverted forms. Jokisch et al. (2005) suggested that the N1 in biological motion processing (which they labeled the N170) reflects a holistic level of processing during which upright figures "pop out" due to their canonical configuration.

Table 1
Summary of previous ERP findings for biological motion perception. All reported between-condition differences were over the posterior parietal–occipital scalp region bilaterally, except as noted. Abbreviations used: PL, point-light; Up, upright; Scr, scrambled; Inv, inverted; LH, left hemisphere; RH, right hemisphere.

Study	Stimuli	Task	N1			N2		
			Up vs. Inv	Up vs. Scr	Inv vs. Scr	Up vs. Inv	Up vs. Scr	Inv vs. Scr
Hirai et al. (2003)	PL walker, L or R direction	Passive viewing	–	Up > Scr RH only 200 ms	–	–	Up > Scr 240 ms	–
Hirai et al. (2005)	PL walker, L or R	Passive viewing	–	Up > Scr 200–300 ms	–	–	Up > Scr 300–500 ms	–
Hirai and Hiraki (2006)	PL walker + masking noise	Human vs. Scr discrimination	–	Up > Scr 200 ms	–	–	Up > Scr 330 ms	–
Jokisch et al. (2005)	PL walker	Human vs. Scr	Up > Inv 150–200 ms	Up > Scr 150–200 ms	No difference	No difference	Up > Scr RH > LH 230–360 ms	Inv > Scr RH > LH 230–360 ms
Krakowski et al. (2011)	Various PL actions	Human vs. Scr	–	Up > Scr RH only 120–170 ms	–	–	Up > Scr LH > RH 190–400 ms	–

They further speculated that this recognition could be based on either perception of static forms from individual video frames, or on some combination of form and motion information. Concerning the N2, Jokisch et al. (2005) suggested that this component may be sensitive to global form and/or a detailed analysis of actions, including recognition of the potential social and psychological implications of human motion. Consistent with this proposal, they provided source localization data for this component consistent with a generator in the STSp. However, Krakowski et al. (2011) also performed source localization on their ERP data in the same time window, and found dipoles that were more posterior than the coordinates of STSp generally reported in fMRI studies of biological motion. Rather, Krakowski and colleagues' findings were more consistent with the coordinates of area V5 which, as noted earlier, overlaps significantly with the EBA (Ferri et al., 2013). Given the degree of uncertainty inherent in ERP source localization, it is difficult to draw strong conclusions from these data. A critical issue remains in that none of the published ERP studies investigating biological motion processing has provided evidence that the effects obtained are specific to biological motion. It may be that the N2 is in fact indexing the recognition of human forms, regardless of whether they move or not.

With this in mind, it is relevant to consider ERP studies of static human forms, of which there are again only a handful. Stekelenburg and De Gelder (2004) showed that static inverted bodies elicited a larger N1 component (adopting the terminology used in the present paper; the authors labeled this component an N170) than upright bodies — similar to the effects for face processing. However, this is apparently true only of bodies with heads; when photos of headless bodies were inverted, the N1 was larger for upright than inverted forms (Minnebusch et al., 2009). Finally, similar to the results for biological motion, scrambled bodies elicited smaller N1s than those in their canonical configuration (Thierry et al., 2006). It is notable that none of these studies of static human forms reported differences in the time window corresponding to the N2 component that has been identified in biological motion studies.

The present study

Thus while the N1 is sensitive to scrambling for both static and moving bodies, we see different patterns of modulations of ERPs for static human forms and biological motion: the N1 to biological motion stimuli is not sensitive to inversion, while the N1 in response to static body images is, but the N2 is sensitive to scrambling of moving but not static bodies. This leads to the hypothesis that the N1 is sensitive to human forms — be they still or moving — and may show sensitivity to inverted static forms because of the fact that these can be recognized more readily than moving point-light forms (which require the integration of information over at least the first several movie frames). The N2 in contrast seems selective for biological motion as it has not been identified in previous studies of static human form perception. However, the type of stimuli, tasks, and stimulus manipulations has varied across studies and critically, no study has directly compared static and moving human forms.

The goal of the present study was to address this gap in the literature, and characterize how ERPs are modulated by both static and moving human forms engaged in a variety of actions, across canonical upright, inverted, and scrambled versions. We included both inverted and scrambled stimuli, both to allow comparison to previous studies, and to assess the effects both of disrupting canonical form (i.e., via inversion) and of eliminating all form information (scrambled motion). Participants first viewed blocks of only static point-light stimuli, followed by stick figures (created by connecting the point-lights of the static stimuli) then motion stimuli. In each block, the stimuli were derived from a set of point-light movies of different human actions (e.g., walking, cycling, and playing tennis; Vanrie and Verfaillie, 2004). Participants were asked to judge if each stimulus was a human form or not; the point-lights in half the stimuli were scrambled to remove the impression of a human form. Stick figures were created by selecting a static frame from biological motion animations and adding lines to connect point-light dots. Each type of stimulus was presented in a separate block, and static form stimuli were presented before the motion stimuli. This was to control for the possibility of expected or implied motion in the stimuli (Kourtzi and Kanwisher, 2000). Participants were not informed that the study was investigating biological motion and did not see any moving stimuli until after the stick figure trials.

On the basis of the literature reviewed above, we hypothesized that static/stick figure stimuli would show inversion and scrambling effects on the N1, while biological motion stimuli would show inversion effects on the N1, and scrambling but not inversion effects on the N2. We further predicted that the scalp distribution of the N1 effects would not differ between static and moving human forms, supporting a model in which human form recognition is supported by the same brain region(s) regardless of whether the form is moving or not. These hypotheses are summarized in Table 2.

Methods

Participants

Twenty-two healthy, naive volunteers (14 females; ages 18–32 years, mean 21 years; years of education from 12 to 19 years, mean 14.9 years) were recruited from the Dalhousie University population and participated in this study. All participants had normal or corrected-to-normal vision and all but one (female) were right-handed. Informed, written consent was obtained prior to participation, and the study was approved by the Research Ethics Board of the Capital District Health Authority. Participants were reimbursed $25 for taking part in the study.

Stimuli

The stimuli used in this experiment were derived from data obtained using motion capture from an actor performing various whole-body actions (Vanrie and Verfaillie, 2004). From a set of 22 different actions, 17 were selected for use in this experiment based on a pilot study conducted using moving point-light sequences to assess ease of recognition. Stimuli that were misidentified by a majority of participants in the pilot study were removed from the set. Text files provided the coordinates for 12 points placed at major joints on the body, and one in the center of the head, as these points moved through space at the rate of 30 Hz. Point-light videos and still images (white dots on a black background) were generated from these using MATLAB (The MathWorks, Inc., Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Inverted and scrambled stimuli were created in MATLAB. Scrambled stimuli were generated by randomizing the location of each point-light in the static images, and the starting location of each point-light in

Table 2
Summary of hypotheses for each stimulus type and contrast. Checkmarks indicate time windows and where significant differences between stimulus types were predicted.

	N1	N2
Static point-light displays		
Upright–scrambled	√	
Inverted–scrambled		
Upright–inverted	√	
Stick figures		
Upright–scrambled	√	
Inverted–scrambled		
Upright–inverted	√	
Biological motion		
Upright–scrambled	√	√
Inverted–scrambled		√
Upright–inverted		√

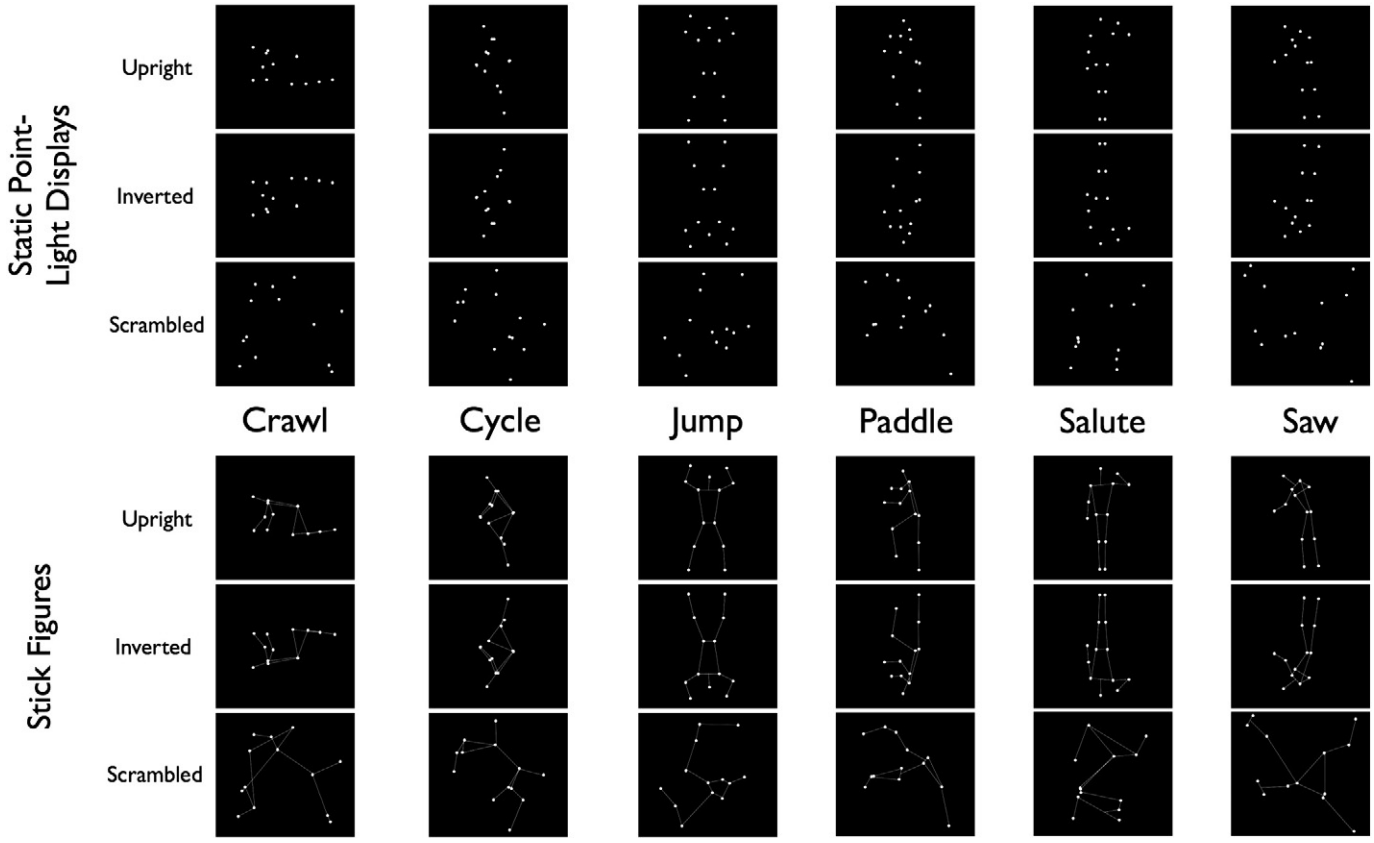


Fig. 1. Example stimuli. Each column shows different versions of one action. Top three rows show the static point-light versions, while bottom rows show the stick figure versions. In each set of three rows, upright, inverted, and scrambled versions are shown.

the motion condition (Bertenthal and Pinto, 1994). All stimuli consisted of square 600 × 600 pixel arrays.

We initially constrained the randomization of point-lights such that the visual angle of each scrambled stimulus matched that of its intact biological counterpart, but found in pilot testing that these stimuli were too difficult to discriminate from biological motion. As such, we relaxed our constraints such that point-lights in all scrambled stimuli could occupy a 400 × 400 pixel region centered in the image. This space corresponded approximately to the maximum height and width of the human form/motion stimuli found in the items.

In the motion condition the motion trajectory of each dot was preserved, thus maintaining all local motion information while disrupting the global percept of human motion. In all conditions scrambled versions were made of both upright and inverted stimuli. Stick figures were made using Adobe Photoshop CS2 (Adobe Systems, Inc., San Jose, CA) to draw thin white lines connecting the point-light dots. In the case of scrambled stimuli, lines were drawn arbitrarily to connect the dots in the scrambled point-light images. Examples of each type of stimulus are shown in Fig. 1.

For the motion block, the scrambled stimuli actually comprised two subtypes: those created from upright movies and those created from inverted movies. In initial analyses we found no reliable differences between these two types of scrambled motion stimuli, so for the purposes of reducing both the complexity of the statistical models and the description of the results, we combined the two types of scrambled stimuli in the analyses reported here.

Procedures

During ERP recording, participants sat in an electrically shielded sound-attenuated booth while viewing stimuli on an LCD monitor from a distance of 110 cm. Participants were asked to remain still and keep their gaze fixed on the center of the screen. Stimuli were presented using DirectRT v2006 (Empirisoft Corporation, New York, NY). There were four blocks of trials comprising three types of point-light stimuli: the first block consisted of static point-light images, the second of stick figures, the third of classic biological motion, and the fourth of the static images again. The second block of static point-light images was included to address a separate research question and those data will be discussed in a separate report. Each block contained all 17 different actions in normal and spatially scrambled arrangements, with both upright and inverted versions of each (68 unique stimuli per block). Each stimulus was presented 3 times per block, for a total of 204 stimuli per block. The stimuli were presented for 1 s each in random order, followed by a screen prompting participants to make a 2-alternative forced-choice response to stimuli: *Human* or *random form*?. Subjects could not respond prior to this prompt. Prior to the start of the experiment, participants were given a few practice trials (using static point-light stimuli only) to ensure that they understood the instructions.

For motion stimuli, this meant looping videos during which the action took less than 1 s (1 video), and cutting off the end of actions that exceeded 1 s (15 videos). The stimuli were specifically designed to loop smoothly (Vanrie and Verfaillie, 2004), so motion discontinuity was not visible for the one looped stimulus. While it is possible that some actions may not have been as recognizable when truncated to 1 s, we only report ERP data from correctly-responded trials, and as the d' analysis in the Results section demonstrates, participants were above chance overall in all conditions. Thus any issues of truncated stimuli should not affect the data we report here.

Behavioral responses were collected using a USB gamepad (Logitech, Fremont, CA). To reduce participant fatigue, self-paced breaks were spaced both throughout and between trial blocks.

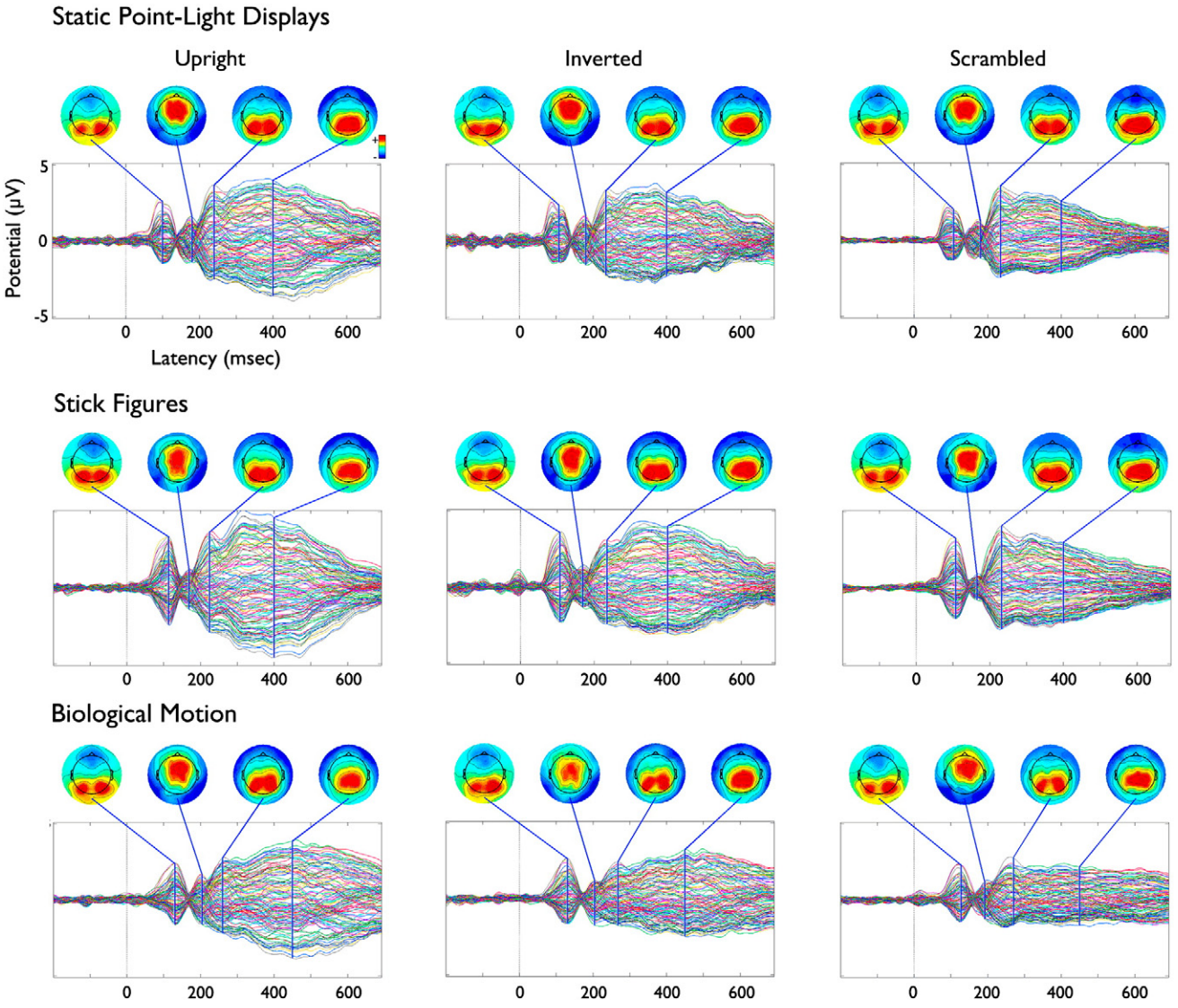


Fig. 2. Butterfly plots showing ERP waveforms from each electrode overlaid on a single axis, for each stimulus type and condition. Positive voltage is plotted up. Scalp voltage topography maps are plotted at time points of local amplitude maxima.

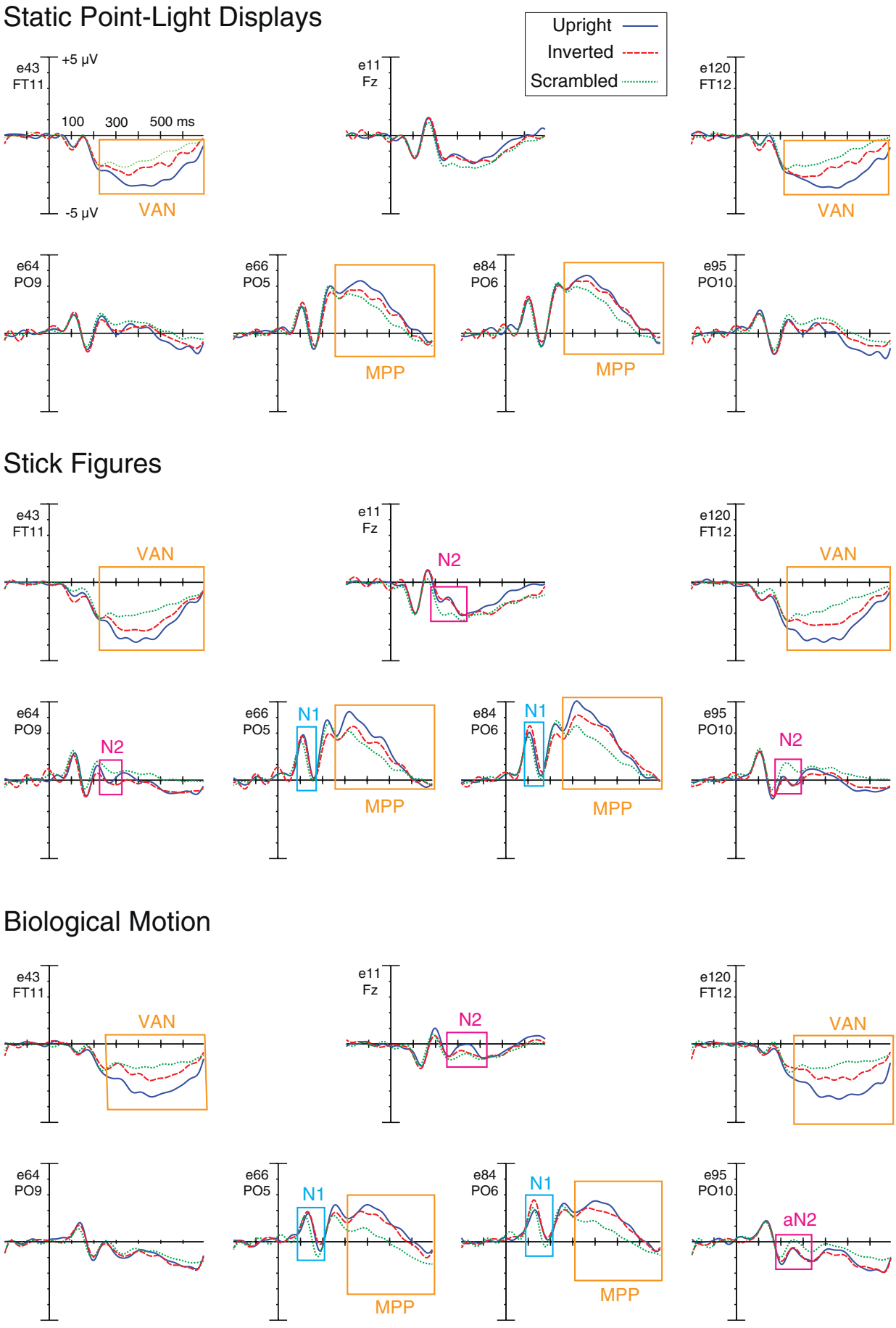
ERP recording and preprocessing

Event-related potentials were recorded from a 128-electrode HydroCel Geodesic Sensor Net (Electrical Geodesics, Inc., Eugene, OR). The electrode montage is shown in Fig. S1 in the Supplementary materials. Data acquisition was done using NetStation software, version 4.3 (Electrical Geodesics, Inc., Eugene, OR). EEG data were digitized with a sampling rate of 256 Hz, using the vertex as reference electrode. EEG data was filtered on-line using a 0.1–100 Hz bandpass filter and stored on a computer for off-line analysis. Post-processing involved several steps. First, using NetStation version 4.3 (Electrical Geodesics, Inc., Eugene, OR), EEG data were filtered off-line using a 0.5–30 Hz bandpass filter, then segmented into epochs consisting of 200 ms pre-stimulus baseline and 750 ms post-stimulus onset, and then baseline corrected by setting the mean of the baseline period to zero for each electrode in each epoch. Trials for which the behavioral response was incorrect were excluded, and subsequent preprocessing steps were conducted in the EEGLAB toolbox version 9.0.8.6 (Delorme and Makeig, 2004) running under MATLAB 2011a (The MathWorks, Inc., Natick, MA). Data were visually inspected and trials with large artifacts (defined as making the EEG difficult or impossible to see, and occurring rarely) were rejected prior to Independent Components Analysis (ICA; Jung et al., 2000). ICA was used to isolate and remove well-characterized artifacts (e.g., blinks, muscular noise) from the data; ICA components were manually rejected on the basis of topographical distribution of activity, power spectra across frequencies, and how consistent and well characterized a given component appeared across experimental trials (Delorme and Makeig, 2004; Jung et al., 2000).

Data analysis

In order to accurately capture the timing and scalp distribution of the differences between conditions for each type of stimulus, we used the mass univariate data approach to statistics described by Groppe et al.

Fig. 3. Grand-averaged ERP waveforms for each stimulus type (upright, inverted, and scrambled), in each block (static point-light displays, stick figures, and biological motion). Electrode positions are given using both their labels on the HydroCel Geodesic Sensor Net, and International 10–10 Equivalents. Positive voltage is plotted up.


Static Point-Light Displays
Upright
Inverted
Scrambled
e43
FT11
+5 μV
-5 μV
100
300
500 ms
VAN
e11
Fz
e120
FT12
VAN
e64
PO9
e66
PO5
MPP
e84
PO6
MPP
e95
PO10
Stick Figures
e43
FT11
VAN
e11
Fz
N2
e120
FT12
VAN
e64
PO9
N2
e66
PO5
N1
MPP
e84
PO6
N1
MPP
e95
PO10
N2
Biological Motion
e43
FT11
VAN
e11
Fz
N2
e120
FT12
VAN
e64
PO9
e66
PO5
N1
MPP
e84
PO6
N1
MPP
e95
PO10
aN2

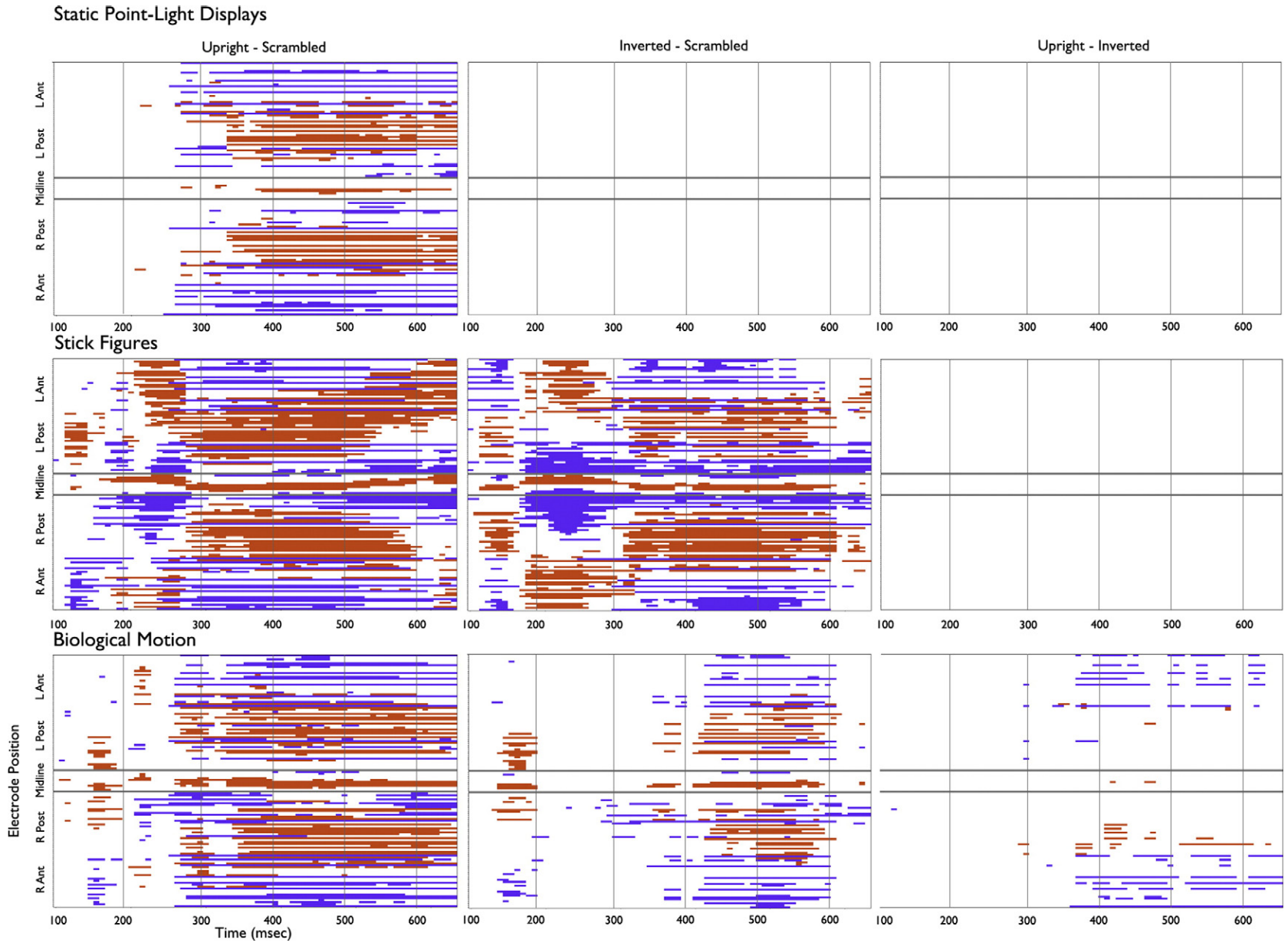


Fig. 4. Raster plots showing the time points and electrodes where significant between-condition *t* scores were found for each stimulus type. The plots have been thresholded at the critical *t* values determined by the FDR method for each contrast. Red indicates positive *t* scores, blue negative *t* scores, and white are failures to reject the null hypothesis.

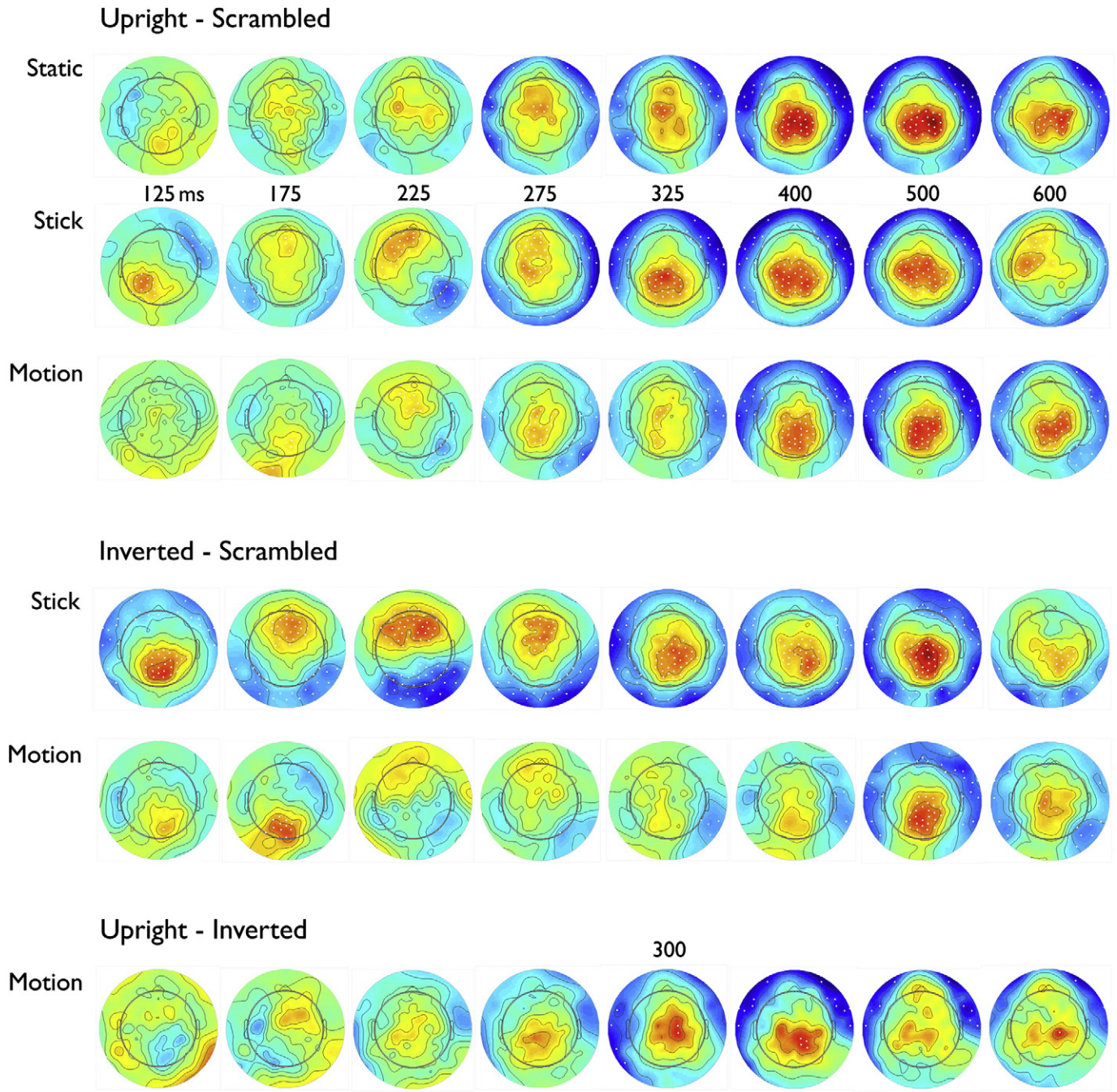


Fig. 5. Statistical parametric maps showing *t* scores testing the null hypothesis of no difference between conditions, for contrasts and stimulus types where significant differences were found. Time points chosen for plotting were selected on the basis of the raster plots (Fig. 4) to emphasize time windows in which statistically significant differences between conditions occurred. The plots have been thresholded at the critical *t* values determined by the FDR method for each contrast, to emphasize the topographical distribution of the effects. Red represents the most positive values, and blue the most negative values. White circles indicate electrodes where the *t* scores were deemed significant by the false discovery rate (FDR) method of multiple comparison correction (Benjamini et al., 2006).

(2011). Data from individual trials were averaged across subjects for each condition and stimulus type. Data were then downsampled to 125 Hz (i.e., 8 ms bins), and then difference waves were computed for each contrast of interest, for each type of stimulus (i.e., upright–scrambled, inverted–scrambled, and upright–inverted, for each of the static point-light, stick figures, and biological motion stimuli). A *t* test was then applied to the data from each electrode, at each time point of each difference wave testing the null hypothesis of zero difference between conditions. These *t* tests were applied to each time point between 100 and 650 ms; this range was selected based on visual inspection of the difference waves. To correct for the large number of statistical tests performed, we applied the two-stage false discovery rate (FDR) method developed by Benjamini et al. (2006). The FDR procedure ensures that the proportion of null hypotheses that are incorrectly rejected is less than the nominal alpha, which in this case was set to .05.

Results

Behavioral data

Accuracy

Mean accuracy rates for each stimulus type and block are given in Table S1. Recognizing the difficulties that arise when proportions are analyzed using an ANOVA (Dixon, 2008), we opted to instead use a

logistic linear mixed effects regression implemented in *R* 3.0.1 (R. Development Core Team, 2013) using the *lmer* function from the *lme4* package (Bates, 2005; Bates et al., 2013). Our model included Block (static, stick, movie) and Condition (upright, inverted, scrambled) as fixed-effects predictors and included random intercepts for subjects and items. Log-likelihood ratio testing was used to evaluate inclusion of each main effect individually as well as the interaction; the full model was supported and all coefficients were significant, $p < .001$.

The main effect of Block was attributable to the fact that in general, participants were less accurate for the initial static block, $z = 5.58$, $p < .001$, relative to the subsequent stick figures, $z = 23.82$, $p < .001$, or movies, $z = 24.49$, $p < .001$. Performance did not differ, however, between the stick and movie blocks, $z = -1.24$, $p = .217$.

The main effect of Condition was modulated by the Block × Condition interaction. Accuracy across all blocks was significantly higher for upright than inverted stimuli, $z = 17.35$, $p < .001$, and also for scrambled than inverted stimuli, $z = 13.83$, $p < .001$. However, the Block × Condition interaction appears to have been driven by a difference in the pattern of accuracy for upright relative to scrambled stimuli across blocks. Specifically, accuracy did not differ between upright and scrambled stimuli for stick figures or biological motion, but did for static point-light displays, with higher accuracy for scrambled stimuli, $z = -12.17$, $p < .001$.

These analyses demonstrated that accuracy was particularly low for inverted, static point-light stimuli, and indeed while accuracy was well above chance (50%) for all other conditions and blocks, it was only 47.9% for these stimuli. This led us to question whether participants had actually been able to reliably distinguish these stimuli from non-biological (scrambled) stimuli in this block. To ensure that participants were sensitive to the biological nature of our stimuli in all blocks and conditions, we computed d' for each. These "sensitivity index" values consider the ratio of correct identifications to false alarms for each condition, in each block, and are shown in Supplementary Table S2. Analyses of these values revealed a pattern comparable to the accuracy data reported above, with d' well above 0 in each case. Importantly, this was true for inverted static figures for which $d' = 1.26$, a value that was significantly greater than the null hypothesis that $d' = 0$, $t(20) = 7.0$, $p < .001$. This finding verifies that while participants had greater difficulty in identifying inverted static point-light forms, they did so at a rate that was significantly higher than that at which they misidentified scrambled stimuli as biological. More generally, the results demonstrate that participants were capable of discriminating the presence of biological motion or form across all conditions.

Reaction times

Reaction time data were analyzed using repeated measures ANOVA in SPSS version 20 (IBM Corporation, Armonk NY, USA). Prior to data analysis, reaction times for incorrect responses as well as those falling outside 3 standard deviations from the mean were excluded, and the data were transformed using the natural logarithm to correct for positive skew. Descriptive statistics for the trimmed data are given in Supplementary Table S3. There was a significant main effect of Block, $F(2,40) = 14.75$, $p < .0001$. There was no main effect of condition, $F(2,40) < 1$, and the Block × Condition interaction also failed to reach statistical significance, $F(4,80) < 1$. Means and standard deviations for each condition, in each block, are shown in Table S2. Post-hoc tests of the main effect of block[3] revealed that responses in the static block were significantly slower than responses in the stick figure block, $t(20) = 5.76$, $p < .0001$ — which is consistent with the idea that the stick figures are intuitively much easier to recognize. People were also faster to recognize stick figures relative to biological motion, $t(20) = 3.65$, $p = .005$, while response times did not differ significantly between the static block and biological motion, $t(20) = 1.49$, $p = .44$. Given that the accuracy analyses showed comparable performance for stick figures and biological motion, the faster reaction times for stick figures than biological motion were not attributable to a speed–accuracy tradeoff.

ERP data

Fig. 2 shows "butterfly" plots for each condition in each time window, showing waveforms from all electrodes and scalp voltage distributions at time points of peak amplitudes. The group-averaged waveforms of the ERPs for each condition are shown at selected electrodes for each stimulus type in Fig. 3. The overall pattern of ERP components was broadly similar across conditions and blocks, and consistent with the prior literature. Examination of electrodes over ventral temporal–occipital scalp regions that have been the focus of prior studies, e.g., PO5, PO6, PO9, and PO10 in Fig. 3, reveals the expected P1–N1–P2–N2 pattern. In the static and stick blocks, the P1, N1, and P2 components (waveform peaks) peaked around 100–110, 170–190, and 220–240 ms, respectively. In the motion block, they peaked between 20 and 50 ms later, at approximately 130, 200, and 270 ms respectively. Looking at the scalp distribution of the waveform peaks in Fig. 2, we see that the P1 peaked bilaterally over occipital sites, followed by the N1 waveform peak maximal over ventral occipito-temporal sites. The P2 had bilateral occipital foci slightly more dorsally than the P1, accompanied by negative potentials (the N2 waveform peak) over ventral temporal and lateral anterior scalp regions (e.g., peaking at around 350 ms at electrode PO10 in Fig. 3).

The results of the mass univariate *t* tests at each electrode, time point, and condition/stimulus type are shown in Fig. 4. Scalp maps showing the distribution of *t* values over time are shown in Fig. 5. For reference, in Fig. 3 we have highlighted time windows on the waveforms where significant effects were obtained, with labels corresponding to those described for each block below. A first observation was that, as we noted in reference to prior studies, the timing of the peak differences between conditions did not correspond to the timing of the waveform peaks observed for the individual conditions in Fig. 2. For example, the earliest differences between conditions occurred between 100 and 200 ms, overlapping with both the P1 and N1 peaks in the waveforms. This was part of our motivation for using the data-driven mass univariate approach to analysis, rather than a more traditional approach in which specific time points and scalp regions are targeted based on observed waveform peaks. A second observation was that, while previous studies have not generally focused on time points after the N2 effect, in the present data it was apparent that the medial posterior negativity and concomitant ventral–lateral negativity scalp topography of the P2 component persisted throughout the remainder of the epoch, with the bilateral occipital foci of the P2 shifting to have a single, more dorsal–medial focus after 400 ms. Hereafter we refer to this effect as the medial parietal positivity/ventral anterior negativity (MPP/VAN).

Static point-light displays

As seen in the top row of Fig. 4, statistically reliable differences for static point-light forms were found only for the upright versus scrambled contrast. Robust significant differences between these conditions began at 250 ms and lasted through the end of the 650 ms epoch. There were no differences between conditions corresponding to the previously described N1 effects for biological motion or stationary static human bodies. Beginning around 250 ms, an enhanced negativity for upright compared to scrambled stimuli was observed over ventral–lateral electrode sites, extending from posterior temporal to frontal–temporal electrodes, and accompanied by a medial–frontal positivity. While similar in timing to the N2 effect previously described for point-light biological motion relative to scrambled motion, the scalp distribution of this effect was rather different in the negativity's having such an anterior extent rather than a

[3] All *p* values were Bonferroni-corrected for number of comparisons unless otherwise noted — 3 comparisons for main effects of block and 3 comparisons for the main effect of condition, and 9 for the Block × Condition interaction.

posterior temporal–occipital focus. After 300 ms the positivity assumed a more medial parietal distribution. This later component is what we have labeled the MPP/VAN.

Stick figures

Compared to the static point-light block, stick figures elicited more extensive differences between conditions. Reliable differences were found between both upright and inverted stick figures relative to scrambled figures, and extended over a greater period of time, as seen in the middle row of Fig. 4.

The first period of significant differences coincided with the timing of the P1 and N1 waveform peaks, from approximately 130 to 200 ms — the time window of the N1 modulation reported in previous studies. We have labeled this the N1 in Fig. 3 to indicate its correspondence in time with the previously-described effect. However, our results did not show the same enhanced posterior temporal negativity for upright relative to scrambled and inverted stimuli that previous studies had reported. Rather, upright and inverted stick figures elicited a stronger positivity over parietal–occipital scalp regions (see electrodes e66 and e84 in Fig. 3) than scrambled figures, and over the anterior scalp a corresponding negativity was larger for human forms than scrambled figures. For upright figures, these effects were restricted to the left posterior scalp (the same electrodes where the P1 waveform peak was maximal) and right anterior regions; for inverted figures the positivity had a midline focus and the corresponding negativity was bilaterally distributed.

Following this enhanced parietal–occipital positivity, from approximately 170 to 275 ms upright and inverted stick figures elicited a stronger negativity than scrambled figures over posterior temporal–occipital electrodes. This effect overlapped with the posterior P2 and N2 waveform peaks, and seems to correspond to the N2 effect described in previous studies. This effect was accompanied by an enhanced positivity for upright and inverted stick figures over medial anterior electrodes. As with the previous time window, the effects for upright stick figures appeared to be lateralized while those for inverted figures did not; in this window however the effect over the posterior scalp was right-lateralized for upright figures, whereas in the earlier window the positivity was left-lateralized.

Following the N2 effect, both upright and inverted stick figures elicited a stronger MPP/VAN than scrambled figures. Significant differences started around 300 ms and lasted through 650 ms. The greater positivity for stick figures than scrambled was focused over the central–posterior midline, and the negativity over anterior ventral–lateral electrodes. This distribution is similar to that observed for static point-light stimuli in the same time period.

Biological motion

As noted earlier, the latency of the early waveform peaks was somewhat later in the biological motion than the stick figure block, by approximately 30 ms. This can be seen in Fig. S2, where waveforms from the three blocks are overlaid, separately for each stimulus type. Thus, although the pattern of experimental effects was generally similar for biological motion as for stick figures, the earliest between-condition differences in the biological motion block occurred approximately 30 ms later than for stick figures. Raster plots of the differences are shown in Fig. 4.

As in the stick figure block, for biological motion the first between-condition effects manifested as an enhanced posterior parietal–occipital positivity and corresponding ventral–lateral anterior negativity for upright and inverted stimuli relative to scrambled ones. This lasted from roughly 150–200 ms, overlapping with both the P1 and N1 waveform peaks, and appearing to correspond to the time window of the N1 effect previously described for biological motion. However, as with the stick figures the effect found in the present study was not the enhanced negativity for upright stimuli relative to other types that was reported in previous studies, and it was largest at parietal–occipital electrodes rather than posterior temporal ones (e.g., electrodes e66 and e84 in Fig. 3, rather than e64 and e95, which are closer to the sites previously reported for the N1 effect).

Subsequent to this effect, upright and inverted biological motion elicited a stronger negativity over posterior temporal–occipital electrode sites — corresponding to the previously-described N2 effect. This again was similar to the pattern observed for stick figures, but for biological motion the negativity was right-lateralized for both upright and inverted stimuli (whereas for stick figures it was right-lateralized only for upright stimuli).

By 400 ms, the MPP/VAN was evident for upright and biological motion stimuli relative to scrambled, and this effect again persisted through the end of the time window analyzed at 650 ms. Qualitatively, the MPP/VAN was significant over more electrodes and for a somewhat more extended time period for upright than inverted biological motions. In addition to the differences between both upright and inverted biological motion and scrambled stimuli — which shared similar timings and scalp distribution with the contrasts between these conditions in the stick figure block — for biological motion we found reliable differences between upright and inverted stimuli.[4] These are shown in the bottom row of Fig. 5. The primary difference between upright and inverted biological motion was a larger MPP/VAN for upright than inverted stimuli, beginning around 300 ms and becoming extensive over the scalp by about 375 ms.[5]

Direct comparison between stick figures and biological motion

The central question of this study was whether biological motion elicits an ERP component or components that are distinct from those elicited by static human forms. On the one hand, qualitatively, the analyses presented thus far do not appear to support such a claim. Both stick figures and biological motion elicited a similar sequence of positive and negative peaks, with similar scalp distributions. The primary differences between the ERPs elicited by static and moving human forms appeared to be in their amplitude and latency, with stick figures eliciting larger-amplitude positivities, and earlier-onset components. On the other hand, the analyses did reveal apparent differences in the sensitivity of these components to our experimental manipulations across blocks. To further explore potential differences between stimulus types, we performed between-block mass univariate analyses.

A concern with directly comparing the same stimulus type (e.g., upright) between blocks is that the stimuli in the two blocks were physically different: between static and motion blocks, the stimuli differed in the absence or presence of motion, while between the stick figure and motion blocks, the stimuli differed both in motion and the absence or presence of lines connecting the point-lights. Such low-level stimulus differences can drive ERP differences that may be attributable to earlier, lower-level stages of processing than were of interest here (Luck, 2005). This is why we had first examined contrasts between well-matched stimuli within each block. Between blocks, an ideal way to compare biological motion with other stimulus types is to compare difference waves. For example, by first deriving the subtraction of upright–scrambled stick figures, and separately upright–scrambled biological motion,

[4] Retrospectively, we observed what seemed to be differences of similar magnitude between these conditions in both the stick figures and static point-light conditions — particularly at the ventral anterior sites. We thus performed a post-hoc analysis at these sites over the approximate time window when between-condition differences were most robust (400–600 ms) and found statistically significant differences between upright and inverted stimuli in all three blocks. Details of these post-hoc analyses are provided in the Supplementary materials.

[5] As noted in the Methods section, the scrambled biological motion stimuli actually comprised both scrambled upright and scrambled inverted biological motion. Although we combined these in a single "scrambled" condition in the analyses reported above, following Troje and Westhoff (2006) we also considered the possibility that the two types of scrambled motion might elicit distinct ERP responses since observers are sensitive to the influence of gravity on local biological motion vectors. We thus conducted a mass univariate analysis comparing the two scrambled conditions. Only one significant difference was found between these conditions, at a single time point at a single electrode (out of a total of over 8700 tests), as shown in Fig. S3. Thus we concluded that there was no meaningful difference between ERPs elicited by upright and inverted scrambled motion.

the ERP effects attributable to consistent low-level stimulus features (e.g., motion) would be subtracted out. We performed these analyses, comparing difference waves for each of the three *contrasts performed previously* (*upright-scrambled*, *inverted-scrambled*, *and upright-inverted*) between biological motion and each other block (i.e., static and stick figures). These well-controlled analyses yielded no significant differences for any of these contrasts.

To further explore possible differences, we additionally performed mass univariate analyses contrasting stimuli from the same condition (i.e., upright, inverted, or scrambled) between biological motion and each other block. Very extensive and robust differences were obtained between blocks for all three types of stimuli. Raster plots of these differences are shown in Fig. S4, and *t* maps in Fig. S5. However, many of these significant differences were due to the increased latency of the early waveform peaks in response to biological motion stimuli that were noted earlier. Thus we observed a significantly larger posterior positivity/anterior negativity pattern around 100–130 ms corresponding to the earlier P1 for static point-light and stick figures, a larger posterior positivity/anterior positivity around 150–170 ms corresponding to the earlier N1, a larger posterior positivity/anterior negativity from 200 to 275 ms corresponding to the earlier P2. The differences were not entirely due to latency, since as can be seen in Fig. S2 the amplitude of the P1 and P2 waveform peaks for both static point-light and stick figures were larger than for biological motion.

Significant differences were also obtained in the MPP/VAN time window, beginning around 300 ms. Because these differences occurred well after the earlier ERP components more closely associated with sensory processing, and because this component shows a much more sustained time course, these differences are less likely to have been driven by physical differences between the stimuli. Upright static point-light displays elicited a significantly larger MPP/VAN than upright biological motion from approximately 300–400 ms, while for inverted stimuli the differences occurred at fewer electrodes, but throughout the 300–650 ms period and for scrambled stimuli, the differences were robust across a large number of electrodes and sustained over the 300–650 ms time window. Stick figures elicited the largest MPP/VAN components for all three stimulus types, with the most widespread differences between conditions from 300 to 450 ms. The significant differences between the scrambled stick figure and biological motion conditions extended over the largest number of electrodes and time points.

In summary, while these analyses confirmed the amplitude and timing differences between blocks, they did not provide any evidence of a distinct ERP component specific to the biological motion condition. These analyses must also be interpreted cautiously, especially in earlier time windows, due to the differences in the physical properties of the stimuli being compared.

Discussion

This study aimed to determine whether ERP components previously reported to be sensitive to biological motion are in fact specific to biological motion, or rather more general to the perception of human beings, be they represented by motion or static form information. To address this question, participants naïve to point-light biological motion stimuli first viewed static point-light displays, followed by stick figures in which the human forms were more easily perceived, followed by biological motion. We compared ERPs elicited by upright, inverted and scrambled versions of each type of stimulus in order to dissociate ERP components sensitive to biological motion from those more generally responsive to the presence of a human form. We hypothesized that if the N2 is a specific index of processing human motion, as previously suggested, it be insensitive to manipulations of static human figures (i.e., inversion and/or scrambling). More generally, we predicted that motionless human forms would show inversion and scrambling effects on the N1 while biological motion stimuli would show only scrambling effects on the N1, but both scrambling and inversion effects on the N2.

Table 3
Summary of results for each stimulus type and contrast. Observed differences that were predicted a priori (see Table 2) are indicated by checkmarks, observed differences that were not predicted are shown as exclamation points, and predicted effects that were not observed are noted with dashes; a checkmark and an exclamation point together indicate that the observed effect was in the opposite direction to that predicted. Results that were significant only in post-hoc testing are noted with asterisks. We have added a column for the MPP/VAN effects that were observed but not predicted in advance.

	N1	N2	MPP/VAN
Static point-light displays			
Upright–scrambled	–	!	!
Inverted–scrambled			!
Upright–inverted	–		!*
Stick figures			
Upright–scrambled	√!	!	!
Inverted–scrambled	!	!	!
Upright–inverted	–		!*
Biological motion			
Upright–scrambled	√!	√	!
Inverted–scrambled	!	√	!
Upright–inverted		–	!

Our hypotheses were only partially borne out, however, as summarized in Table 3. While we did observe differences in N1 amplitude between scrambled and upright stimuli for both stick figures and biological motion, these effects went in the opposite direction from those reported in previous biological motion ERP studies, and we did not N1 amplitude differences between upright and inverted biological motion. There was evidence of sensitivity of the N1 to inverted human forms, however, as these elicited stronger negative potentials in the N1 time window for inverted than scrambled stimuli, both for stick figures (as predicted) and for biological motion (which had not been found in one previous study that examined this contrast).

The most significant finding of this study was that while previous ERP studies of biological motion had suggested that the N2 is specifically sensitive to biological motion, in fact the N2 showed identical sensitivity to biological motion and stationary human forms. We replicated previous studies in finding stronger negativities for both upright and inverted biological motion than scrambled stimuli, but we also found these effects for stick figures and for upright static point-light figures. This finding suggests that the N2 is not specifically sensitive to biological motion, as previously suggested, but rather may index the recognition of human forms more generally. An additional, important finding of this study was the identification of a previously-unreported component that is sensitive to both scrambling and inversion of human form stimuli — both moving and stationary. Consistent across all stimulus types in this study was a late (300–600 ms) posterior medial positivity accompanied by a ventral–lateral anterior negativity, which we have labeled the MPP/VAN. This component was largest in amplitude for upright stimuli, and smallest for scrambled stimuli.

More generally, no component was identified that was specific to biological motion — the sequence and scalp distribution of the potentials evoked by static and moving human forms, and their general sensitivity sensitivity to stimulus manipulations was highly similar across stimulus types. This is consistent with models in which a network of brain regions, including those sensitive to form and motion information, work in concert to support the recognition of static and moving human forms (e.g., Lange and Lappe, 2006). On the other hand, there were some differences in how the ERPs evoked by static and moving stimuli were influenced by our stimulus manipulations. Most notably, the only difference between upright and inverted stimuli found in the mass univariate analyses was for the MPP/VAN, for biological motion. However, this must be qualified by the fact that post-hoc testing

found similar effects for stick figures and static point-light images, as well as the fact that previous studies had found inversion effects for both biological motion (Jokisch et al., 2005) and stationary figures (Minnebusch et al., 2009; Stekelenburg and De Gelder, 2004). Another difference was that while the N2 differences between upright and scrambled stimuli were right-lateralized for both stick figures and biological motion, for the inverted–scrambled contrast this was only true for biological motion. The right-lateralization is consistent with activation patterns observed for both STSp and EBA in fMRI studies (Downing et al., 2001; Grossman et al., 2000), however the present results suggest that inverted, static human forms evoke more bilateral activity whereas inverted biological motion does not.

Finally, the amplitude and timing of the waveform peaks differed between stimulus types. Waveform peaks for biological motion were consistently about 30 ms later than for the other stimulus types, although this may simply reflect the fact that motion necessarily takes place over multiple frames, and 30 ms corresponds roughly to the time required to present 2 subsequent frames at the refresh rate of our monitor (60 Hz). The largest P1, P2, and MPP/VAN components occurred for stick figures and the largest N1 for static point-light figures. While differences in the earlier components may be attributable to physical differences between the stimuli, the differences in MPP/VAN amplitude may be more directly associated with cognitive differences in how the stimuli were processed. These differences require replication however, as while they were observable qualitatively, we did not find any statistically significant differences when stimulus manipulation contrasts (e.g., upright–scrambled) were compared between stimulus types (e.g., stick figures vs. biological motion).

In what follows we relate our findings to the previous literature, with reference to the N1 and N2 effects previously described, as well as our newly-discovered MPP/VAN effects.

N1

The earliest ERP effects that have been associated with biological motion perception typically occurred from 150 to 200 ms — the effect that we label the N1. In previous studies, the N1 effect manifested as a larger negativity for upright than scrambled stimuli, both for biological motion (Hirai et al., 2003, 2005; Jokisch et al., 2005; Krakowski et al., 2011) and static human forms (Thierry et al., 2006). The N1 had also shown sensitivity to inversion, with larger negativities observed to upright than inverted biological motion (Jokisch et al., 2005) and (headless) bodies (Minnebusch et al., 2009). Surprisingly, while we found significant differences between conditions in this time window for both stick figures and biological motion, these went in the opposite direction to previous studies, were maximal over parietal–occipital rather than more ventral, posterior temporal electrodes, and showed different sensitivity to stimulus manipulations. Specifically, we found significantly larger negativities in this time window to scrambled stimuli than to either upright or inverted human forms, and we did not find any differences between upright and inverted stimuli.

In trying to understand the source of this fundamental difference between our and previous studies, the most striking, consistent difference is the selection of stimuli. In most previous studies, the biological motion stimuli were exclusively point-light walkers. In some studies (Hirai and Hiraki, 2006; Hirai et al., 2003, 2005) these were a single, artificially-generated stimulus that moved either rightward or leftward, while in another study they were motion-capture-derived sequences from 40 different people (Jokisch et al., 2005). In only one study was any variety of human actions used (Krakowski et al., 2011), and even in that study only 10 distinct actions were used, each repeated 30–35 times over the course of the study. In contrast, we used 17 distinct actions, each repeated only 3 times per block. Thus each subject saw a given action performed as biological motion only 3 times in our study, as opposed to 30–60 times in previous studies. The significantly greater variety in actions to be recognized in our study may well have been the critical factor driving the differences from previous findings.

Jokisch et al. (2005) proposed that the N1 effect reflects the recognition of a human figure, and more specifically, an early stage at which bodies in their canonical orientation are processed holistically and "pop out". The timing of these effects is consistent with the findings of Cutting et al. (1988), who showed that 200 ms is sufficient to recognize point-light biological motion. Jokisch suggested that the larger N1 to upright than either inverted or scrambled stimuli in their study reflected the fact that the latter stimulus types took longer to recognize due to their non-canonical configurations. In the present study, the recognition process may have been qualitatively different from previous studies. That is, rather than simply discriminating a point-light walker from a scrambled stimulus (or one of a limited set of frequently-repeated stimuli) subjects in the present study were challenged to recognize human forms performing a wide variety of actions (including some in which the actor was not in a canonical standing position but rather sitting or bending over), each of which was seen very few times. Thus while in previous studies subjects could rely on a holistic mental template of a human form in a particular orientation for rapid, holistic identification, in our study more detailed analysis was likely required that may have precluded rapid "pop out" recognition of upright forms and the consequent N1 enhancement for upright stimuli.

Given that in our study, both the scalp distribution and response properties of this early effect differed from previous studies, it is unclear whether it should even be considered the same "N1" effect at all. The effect we observed overlaps with both the N1 waveform peak and the preceding P1 waveform peak, and the scalp distribution of the between-condition difference is consistent with that of the P1 waveform peak (see the first scalp map in each of the panels of Fig. 2). The N1 enhancements that we observed for scrambled relative to other stimulus types may have been due to the fact that scrambled stimuli "popped out" relative to human stimuli because the dots were, on average, spread over a wider range (though they were restricted to the overall ranges of the human stimuli).

At any rate, the fact that the existence and direction of this early effect differs across studies — both for static and moving human forms — suggests that while this effect is sensitive to differences between human forms and scrambled versions, it does not specifically index either biological motion perception, nor even human form perception more generally. Effects in this time window may instead be a more generic reflection of the visual system's sensitivity to task-related manipulations in object recognition tasks.

N2

As in previous studies of biological motion, we observed an enhanced negativity over posterior temporal electrode sites for both upright and inverted biological motion relative to scrambled motion. However, while previous studies had associated this effect with biological motion perception, we found the same modulation for stationary stick figures. Our data also extend previous findings by documenting the anterior midline positivity that accompanies the N2.

The two previous studies that attempted to localize the sources of the N2 for biological motion stimuli both highlighted the role of the posterior temporal lobe, though Jokisch et al. (2005) implicated the STSp while Krakowski et al. (2011) found a more posterior focus, consistent with EBA and/or V5. Sources in these regions are consistent with fMRI studies of both biological motion and static body perception, and the low spatial resolution and inherent uncertainty of ERP source localization may limit our ability to discriminate between these nearby sources. Furthermore, while STSp has been the focus of many studies of biological motion, the EBA and FBA also show selective activation to biological motion (Jastorff et al., 2009; Peelen et al., 2006). Indeed, Jastorff et al. (2009) showed that EBA and FBA integrate form and motion cues in human form and action perception, and "link the portrayed action

with the body" (p. 7327), while STSp provides more detailed analysis of human actions, including identifying task-relevant features. We suggest that the N2 reflects the processes associated with the EBA/FBA, of integrating form and (when available) action information.

Medial posterior positivity/ventral–lateral negativity (MPP/VAN)

While not a focus of previous studies, in the present study the MPP/VAN was the most robust and sensitive ERP index of human form and action perception. The effect was significant across a large number of electrodes and for an extended period of time, exceeding 300 ms. It was also the only effect that discriminated upright from inverted stimuli, and did so most reliably for biological motion stimuli though post hoc testing showed similar effects for stationary point-light and stick figures. The MPP/VAN was consistently highest in amplitude for upright human forms, smallest for scrambled, and intermediate for inverted human forms.

While this effect has not previously been given much attention, some studies included scalp topography maps extending into the time window in which we found the MPP/VAN. These do consistently show a medial parietal positivity along with an anterior/ventral negativity for upright biological motion relative to scrambled motion, from 400 to 500 ms (Hirai and Hiraki, 2006; Krakowski et al., 2011). Krakowski et al. (2011) performed a timepoint-by-timepoint analysis up to 500 ms and found significantly larger positivity for upright than scrambled biological motion, but only when the stimuli were attended and task-relevant. Our results extend these previous findings by demonstrating that this effect is (a) also larger for upright than inverted stimuli, and for inverted relative to scrambled; (b) paired with a VAN, and (c) elicited by both biological motion and static body forms.

Krakowski et al. (2011) interpreted the MPP they observed as "cognitive processes involved in decoding the meaning of the activity displayed by the motion stimulus" (p. 381). Our data indicate that this effect is not limited to biological motion — similar processes are engaged when viewing static images of humans performing actions. Biological motion may however make these actions more clear and salient, which may be why the upright–inverted difference for the MPP/VAN was only statistically significant in the biological motion condition.

Given that earlier components also distinguished human forms from scrambled stimuli, it is reasonable to speculate that the MPP/VAN does not reflect the earliest stages of recognizing human forms, but rather more complex processing. In agreement with Krakowski and colleagues, we believe that the notion of deriving "meaning" from the stimuli is a reasonable hypothesis as to the function of the processing reflected by this component. The response pattern of the MPP/VAN — showing the strongest responses to depictions of humans performing actions in their canonical orientations, and discriminating these from scrambled stimuli only in attended, task-relevant conditions — is strikingly similar to Jastorff and Orban (2009)'s characterization of the role of the STSp: performing detailed analysis of actions that are relevant to the viewer. We thus suggest that the MPP/VAN reflects the activation of the STSp (and possibly other related brain areas) in recognizing (or attempting to recognize) what a human is doing. Importantly, this information may be derived from either motion or static images of human bodies.

This interpretation of the MPP/VAN is consistent with our findings for the static point-light block. We did not find any significant N1 effects for these stimuli, and in the N2 time window the negativity showed a more anterior distribution along with a medial positivity that was more similar to the later MPP/VAN than to the N2 effects observed for stick figures or biological motion in this or previous studies. At the same time, we included in the analysis only trials on which participants correctly discriminated biological from non-biological forms, and the d' analysis demonstrated that subjects discriminated these stimuli at rates better than chance. We thus suggest that the static point-light condition elicited only a MPP/VAN — a fact which suggests that this, rather than the N2, is the ERP effect that reflects conscious recognition of humans performing actions and underlies the successful discrimination of human forms from scrambled stimuli. When additional information is available, the earlier stages indexed by the N2 — or even the N1, under conditions where discrimination is facilitated by a narrow repertoire of possible stimuli such as a single point-light walker vs. scrambled motion — may show differential responsively. However, these differential responses do not appear obligatory for successful discrimination of human from non-biological actions, while the MPP/VAN does.

Conclusion

Our data suggest that while ERPs are sensitive both to the presence of a human form and its inversion, there is no ERP component that is specific to biological motion perception — in contrast to previous claims regarding the N2. Our conclusions are consistent with models which posit that human actions are recognized through distributed activation across this network of regions, and that the perception of human form and motion information are tightly integrated. Our data suggest that the N2 effect reflects the integration of form and motion information, possibly supported by brain regions including EBA and FBA, and therefore that human form and motion information are integrated by about 200 ms after stimulus onset. We further propose that the later MPP/VAN effect reflects efforts to recognize the activities engaged in by human forms, likely supported by the STSp and possibly other brain areas.

Acknowledgments

We are grateful to the following people for their help in this study: J. Bray, D. Groppe, H. Matheson, and P. McMullen. A.J.N. was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) and by the Canada Research Chairs Program; J.M.F. was supported by an NSERC Postgraduate Scholarship; and N.C.W. was supported by an NSERC Summer Undergraduate Research Award.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.09.026.

References

Bates, D.M., 2005. Fitting linear mixed models in R. R News 5, 27–30.
Bates, D., Maechler, M., Bolker, B., 2013. lme4: Linear Mixed-effects Models Using S4 Classes.
Beauchamp, M.S., Lee, K.E., Haxby, J.V., Martin, A., 2002. Parallel visual motion processing streams for manipulable objects and human movements. Response 34, 149–159.
Beintema, J.A., Lappe, M., 2002. Perception of biological motion without local image motion. Proc. Natl. Acad. Sci. U. S. A. 99, 5661–5663.
Beintema, J.A., Georg, K., Lappe, M., 2006. Perception of biological motion from limited-lifetime stimuli. Percept. Psychophys. 68, 613–624.
Benjamini, Y., Krieger, A.M., Yekutieli, D., 2006. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93, 491–507.
Bertenthal, B.I., Pinto, J., 1994. Global processing of biological motions. Psychol. Sci. 5, 221–224.
Blake, R., Shiffrar, M., 2007. Perception of human motion. Annu. Rev. Psychol. 58, 47–73.
Bonda, E., Petrides, M., Ostry, D., Evans, A., 1996. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. J. Neurosci. 16, 3737–3744.
Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436.
Cutting, J.E., Moore, C., Morrison, R., 1988. Masking the motions of human gait. Percept. Psychophys. 44, 339–347.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.
Dixon, P., 2008. Models of accuracy in repeated-measures designs. J. Mem. Lang. 59, 447–456.
Downing, P.E., Jiang, Y., Shuman, M., Kanwisher, N., 2001. A cortical area selective for visual processing of the human body. Science 293, 2470–2473.

Ferri, S., Kolster, H., Jastorff, J., Orban, G.a, 2013. The overlap of the EBA and the MT/V5 cluster. NeuroImage 66, 412–425.
Garcia, J.O., Grossman, E.D., 2008. Necessary but not sufficient: motion perception is required for perceiving biological motion. Vis. Res. 48, 1144–1149.
Groppe, D.M., Urbach, T.P., Kutas, M., 2011. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. Psychophysiology 48, 1711–1725.
Grossman, E.D., Blake, R., 2002. Brain areas active during visual perception of biological motion. Neuron 35, 1167–1175.
Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., Blake, R., 2000. Brain areas involved in perception of biological motion. J. Cogn. Neurosci. 12, 711–720.
Hirai, M., Hiraki, K., 2006. The relative importance of spatial versus temporal structure in the perception of biological motion: an event-related potential study. Cognition 99, B15–B29.
Hirai, M., Fukushima, H., Hiraki, K., 2003. An event-related potentials study of biological motion perception in humans. Neurosci. Lett. 344, 41–44.
Hirai, M., Senju, A., Fukushima, H., Hiraki, K., 2005. Active processing of biological motion perception: an ERP study. Cogn. Brain Res. 23, 387–396.
Jastorff, J., Orban, G.A., 2009. Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing. J. Neurosci. 29, 7315–7329.
Jastorff, J., Kourtzi, Z., Giese, M.A., 2009. Visual learning shapes the processing of complex movement stimuli in the human brain. J. Neurosci. 29, 14026–14038.
Johansson, G., 1973. Visual perception of biological motion and a model for its analysis. Percept. Psychophys. 14, 201–211.
Jokisch, D., Daum, I., Suchan, B., Troje, N.F., 2005. Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis. Behav. Brain Res. 157, 195–204.
Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111, 1745–1758.
Kourtzi, Z., Kanwisher, N., 2000. Activation in human MT/MST by static images with implied motion. J. Cogn. Neurosci. 12, 48–55.
Krakowski, A.I., Ross, L.A., Snyder, A.C., Sehatpour, P., Kelly, S.P., Foxe, J.J., 2011. The neurophysiology of human biological motion processing: a high-density electrical mapping study. NeuroImage 56, 373–383.
Lange, J., Lappe, M., 2006. A model of biological motion perception from configural form cues. J. Neurosci. 26, 2894–2906.
Luck, S.J., 2005. An Introduction to the Event-related Potential Technique. MIT Press, Cambridge MA US.
Matheson, H.E., Mcmullen, P.A., 2010. Neuropsychological dissociations between motion and form perception suggest functional organization in extrastriate cortical regions in the human brain. Brain Cogn. 74, 160–168.
Michels, L., Lappe, M., Vaina, L.M., 2005. Visual areas involved in the perception of human movement from dynamic form analysis. NeuroReport 16, 1037–1041.
Minnebusch, D.A., Suchan, B., Daum, I., 2009. Losing your head: behavioral and electrophysiological effects of body inversion. J. Cogn. Neurosci. 21, 865–874.
Peelen, M.V., Wiggett, A.J., Downing, P.E., 2006. Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. Neuron 49, 815–822.
Pelli, D.G., 1997. The Video Toolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10, 437–442.
R. Development Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Saygin, A.P., 2007. Superior temporal and premotor brain areas necessary for biological motion perception. Brain 130, 2452–2461.
Servos, P., Osu, R., Santi, A., Kawato, M., 2002. The neural substrates of biological motion perception: an fMRI study. Perception 12, 772–782.
Stekelenburg, J.J., De Gelder, B., 2004. The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. Cogn. Neurosci. Neuropsychol. 15, 9–12.
Thierry, G., Pegna, A.J., Dodds, C., Roberts, M., Basan, S., Downing, P., 2006. An event-related potential component sensitive to images of the human body. NeuroImage 32, 871–879.
Troje, N.F., Westhoff, C., 2006. The inversion effect in biological motion perception: evidence for a "life detector"? Curr. Biol. 16, 821–824.
Vaina, L.M., Gross, C.G., 2004. Perceptual deficits in patients with impaired recognition of biological motion after temporal lobe lesions. Proc. Natl. Acad. Sci. U. S. A. 101, 16947–16951.
Vaina, L.M., Lemay, M., Bienfang, D.C., Choi, A.Y., Nakayama, K., 1990. Intact biological motion and structure from motion perception in a patient with impaired motion mechanisms: a case study. Vis. Neurosci. 5, 353–369.
Vaina, L.M., Solomon, J., Chowdhury, S., Sinha, P., Belliveau, J.W., 2001. Functional neuroanatomy of biological motion perception in humans. Proc. Natl. Acad. Sci. U. S. A. 98, 11656–11661.
Vanrie, J., Verfaillie, K., 2004. Perception of biological motion: a stimulus set of human point-light actions. Behav. Res. Methods 36, 625–629.