

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/338441455>

Features and feelings. Objective and subjective features critical for emotional recognition in whole-body movements

Preprint · January 2020

DOI: 10.31234/osf.io/mr9tz

CITATION

1

READS

327

3 authors:



Marta Poyo Solanas

Maastricht University

11 PUBLICATIONS 59 CITATIONS

[SEE PROFILE](#)



Maarten Vaessen

Maastricht University

36 PUBLICATIONS 1,141 CITATIONS

[SEE PROFILE](#)



Beatrice de Gelder

Maastricht University

301 PUBLICATIONS 15,683 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



face processing in ASD [View project](#)



Non-conscious processing of bodies under CFS [View project](#)

Features and feelings. Objective and subjective features critical for emotional recognition in whole-body movements

Marta Poyo Solanas¹, Maarten Vaessen¹, & Beatrice de Gelder^{1,2,*}

¹ Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Limburg 6200 MD, The Netherlands.

² Department of Computer Science, University College London, London WC1E 6BT, UK.

* Address correspondence to Beatrice de Gelder, Department of Cognitive Neuroscience, Brain and Emotion Laboratory, Faculty of Psychology and Neuroscience, Maastricht University, Oxfordlaan 55, 6229 EV Maastricht, The Netherlands. Tel. +31 433881437. E-mail: b.degelder@maastrichtuniversity.nl

Running title: body features from affective movement

Keywords: Body, Emotion, Movement-features, Perception, Posture

Abstract

Humans are experts at recognizing intent and emotion from other people's body movements; however, the underlying mechanisms are poorly understood. Here, we investigated the role of kinematic and postural information in affective whole-body movement perception. For this purpose, quantitative features of body posture and kinematics were related to emotional categories as well as to behavioural ratings of feature descriptors. Overall, postural rather than kinematic cues seemed to be more relevant in the discrimination between emotional movements, both for the computed and the behavioural features. Particularly, limb angles and symmetry appeared to be the most relevant ones. These findings were observed independently of whether or not the time-related information was preserved in the computed features. Interestingly, behavioural ratings showed a clearer distinction between affective movements than the computed counterparts. Finally, the perceived degree of approach/avoidance behaviour seemed to be especially critical for the recognition of fear.

Introduction

Social species spend considerable time watching each others' body postures and movements since this information is highly relevant for their own behaviour. Often, our understanding of the actions and emotions expressed by the body is direct and automatic as, for example, when we react to an aggressive posture by stepping back. Other times, we are consciously aware of the feelings triggered by the body postures and movements (e.g. fear, anger, dominance) (de Gelder, 2006; de Gelder et al., 2010). However, neither the body features driving the perception of the emotional content nor the features that play a prominent role in our conscious feelings have been systematically studied. A better understanding of the core features of nonverbal communication will have a crucial impact on our understanding of affect and social interaction, and will directly benefit many areas of society, especially health care, where this knowledge could be useful in the treatment of affective communication disorders.

Most studies have so far investigated how body movements convey emotion by relating verbal descriptions of posture and movement properties to qualitative emotion categories (De Meijer, 1989; Wallbott, 1998). For example, important postural features in discriminating between affective states have been found including head inclination, which is typical for sadness, or limb flexion, which observers associate with the expression of anger (Coulson, 2004; Wallbott, 1998). Other candidates are the degree of lateral opening of the body (e.g. the body is more extended for happy than for sad), the vertical extension of the body (e.g. hands are often raised for happy but remain low for sad), symmetry (e.g. joy is often depicted by symmetric up and down hand movement) or the directionality of the movement (e.g. forward whole-body movement depicts anger) (for a review see Kleinsmith & Bianchi-Berthouze, 2012).

In contrast to the use of qualitative descriptions and categories, computer scientists are increasingly interested in modelling the properties of body postures and movements (Kleinsmith & Bianchi-Berthouze, 2012; Niewiadomski et al., 2017; Patwardhan, 2017; Piana, Stagliano, Odone, Verri, & Camurri, 2014; Roether, Omlor, Christensen, & Giese, 2009; Vaessen, Abassi, Mancini, Camurri, & de Gelder, 2018). This requires a detailed analysis of the different types of information conveyed by body movements: kinematic (e.g. velocity), dynamics (e.g. mass and force) and posture/form information and its changes over time (Atkinson, Tunstall, & Dittrich, 2007). With regard to kinematics, it has been found that velocity, acceleration, and jerkiness strongly influence the perception of emotion in expressive arm (Paterson, Pollick, & Sanford, 2001; Pollick, Paterson, Bruderlin, & Sanford, 2001; Sawada, Suda, & Ishii, 2003) and also in whole-body movements (Roether et al., 2009).

However, the majority of studies investigating the contribution of form and motion information to emotional attribution employed point-light displays (PLD) (Johansson, 1973). While the use of PLDs allows for the control of possible confounds in emotional recognition, such as identity and gender, as well as permitting systematic variations of kinematic and postural features, they are far from representing natural stimuli. Dance movements have also been used (Camurri et al., 2016; Vaessen et al., 2018) but, although more naturalistic, they are often exaggerated and do not represent day-to-day emotional movements, actions or social interactions. The use of static body pictures, on the other side, obviates the dynamic nature of affective body signals. Therefore, more naturalistic dynamic stimuli are needed to gain insight on how low-level visual body attributes contribute to the perception of specific affective states. Their use, however, comes with difficulties since the configuration of whole-body expressions presents a high dimensionality, and its overall shape varies strongly during movement (Roether et al., 2009; Schindler, Van Gool, & de Gelder, 2008).

The goal of the present study was pioneer a quantitative representation of naturalistic whole-body movements using computational features. Particularly, we investigated the role of kinematic and postural information in whole-body movement perception. For this purpose, quantitative features were derived from the position of the actors' main joints, obtained with the state-of-the-art 2D pose estimation library OpenPose (Cao, Simon, Wei, & Sheikh, 2017). Furthermore, we also obtained behavioural features derived from measures of subjective perception. This allowed us to investigate the relationship between the quantitatively computed features and emotion categories as well as with the subjective ratings of feature descriptions.

Materials and methods

Participants

Thirty-two volunteers participated in the behavioural experiment, but only the data of thirty (mean age = 22.97; age range = 19-36; ten males; four left-handed participants, one of them male) were included in the analysis due to technical issues in data recording. All participants had normal or corrected-to-normal vision and a medical history without any psychiatric or neurological disorders. The experiment was performed in accordance with the Declaration of Helsinki and all procedures followed the regulations of the Ethical Committee at Maastricht University. All participants provided written informed consent before taking part in the experiment. Participants either received credit points or were reimbursed with vouchers.

Stimuli

Stimuli consisted of 56 one-second video clips (25 frames) of whole-body movements. In each video, a male actor expressed one out of three possible emotional body movements: happy, fearful or angry. The stimulus set also included neutral body actions such as coughing, pulling

the nose or walking. Therefore, this experiment consisted of four categories (i.e. happy, fear, anger, neutral), each of them consisting of 14 videos with seven male actor identities.

The stimuli were computer-edited using Ulead, After Effects and Lightworks (EditShare). To avoid triggering facial perception processes, the faces of the actors were blurred with a Gaussian mask so only the information of the body was available. In addition, all actors were dressed in black and filmed against a green background under controlled lighting conditions. The video clips included in this experiment belonged to a larger stimulus set and were selected based on a high recognition accuracy (>80%). For more information regarding the recording and validation of these stimuli, see Kret, Pichon, Grèzes, & de Gelder (2011).

Pose estimation

A state-of-the-art 2D skeleton extraction library called OpenPose (v1.0.1) (Cao et al., 2017) was used to infer each actor's pose in the video stimuli. OpenPose uses a convolutional neural network to estimate the position of the main joints in a total of 18 *keypoints* (i.e. ears, eyes, nose, neck, shoulders, elbows, hands, left and right part of the hip, knees and feet). Each *keypoint* is defined by its x and y image coordinates and a confidence value indicative of the algorithm's certainty in determining the position of the joint. Subsequently, OpenPose uses part affinity fields to associate the *keypoints* in order to produce an anatomically inspired skeleton (see **Fig. SM1** Supplementary Materials). Due to the blurring of the face in our video clips, the estimation of the location of eyes and ears was often inaccurate. These *keypoints* were disregarded for further analysis giving that our purpose is the computation of kinematic and postural body features. However, the *keypoint* corresponding to the nose was kept for further analysis as a reference for the position of the head. Thus, three values were obtained for the remaining 14 *keypoints* and 25 frames for each of the 56 videos of the stimulus set. In addition, visual inspection of the estimated joint positions was performed to assess the accuracy of the

algorithm and manual corrections were performed when necessary with the help of Adobe Photoshop CS6's coordinate system (v13.0, Adobe Systems Inc., San Jose, CA, USA).

Feature definition

To investigate the possible contribution of kinematic and postural body attributes to the processing and recognition of emotional movements, several quantitative features were computed giving their importance in previous work (for a review see Kleinsmith & Bianchi-Berthouze, 2012). These features include velocity, acceleration, vertical movement, symmetry, limb angles and three different computations of body contraction (i.e. shoulder ratio, surface and limb contraction). These features were calculated using custom code in MATLAB (vR2017a, The MathWorks Inc., Natick, MA, USA) from the x- and y-coordinates of the 14 *keypoints* and 25 frames of each of the 56 videos. Although each feature was calculated within each frame, the time information was later averaged (see Supplementary Materials for more information on feature definition).

Experimental design, task and procedure

For this behavioural experiment, the 56 videos that comprised the body-movement stimulus set were presented in four runs lasting approximately 15 minutes, respectively. In each run, 14 videos were shown. Each trial consisted of 100ms fixation period followed by one-second video presentation. Immediately after the video, participants were required to answer 11 questions with regard to kinematic- (i.e. amount of movement, fast movement, vertical movement, direction of the movement), posture- (i.e. body contraction, symmetry), emotion- (i.e. emotional category, intensity, familiarity, valence), and action- (i.e. action category) related information of the body movement displayed in the video (see Supplementary Materials). The questions about kinematic and postural aspects were closely related to the computed features, and were rated on a seven-point scale using a computer mouse. The emotional and action

categorization questions required a forced-choice answer (see Supplementary Materials). Before the actual experiment, instructions and a practice run were provided to the participants. The stimulus presentation order was randomized, both within and between runs, for each participant. The stimuli were displayed using PsychoPy2 (v1.90.0) (Peirce, 2007, 2009) in the centre of a computer screen (screen resolution = 1920 x 1200; screen refresh rate = 60Hz) under controlled lighting conditions. The stimuli spanned 14.03 degrees of visual angle.

Representational similarity analysis

Relations among the computed features and behavioural ratings were calculated by means of representational similarity analyses (Kriegeskorte et al., 2008; Nili et al., 2014) in MATLAB (vR2017a, The MathWorks Inc., Natick, MA, USA). This approach involves the comparison of pairs of stimuli values to determine their representational dissimilarity. RSA characterizes this representation by means of representational dissimilarity matrices (RDM), which are symmetrical. The diagonal entries reflect comparisons between identical stimuli and were defined as zero. Each off-diagonal value indicates the dissimilarity between values associated with two different videos.

Computed-feature and behaviour-based RDMs. Based on the computed features for each of the 56 videos, RDMs were constructed by defining a dissimilarity value for all stimulus pairs in Euclidean distance (i.e. first-level feature RSA). For the emotional categories RDM, dummy variables were used such that the same emotion had zero dissimilarity with itself while two different emotions presented a dissimilarity of $\sqrt{2}$. A first-level RSA was also performed with the behavioural ratings using Euclidean distance. For each rating, a group RDM was produced by averaging the RDMs of all participants. For the emotional and action ratings, dummy variables were used as in the case of the emotional category RDM. This analysis generated 56x56 distance matrices for both the computed features and the behavioural ratings. To examine

the relationship between each computed feature and perceptual rating, Spearman's rank correlations were carried out, resulting in second-level RDMs. Spearman's correlations were also performed to investigate possible correlations between features and behavioural ratings, respectively.

Classification regression trees

In order to investigate the relative importance of the kinematic and postural features as well as the participant's ratings in the classification of affective body movements, decision tree classifiers (Loh, 2002) were implemented in the Machine Learning Toolbox (v11.1) from MATLAB (vR2017a, The MathWorks Inc., Natick, MA, USA). The classification was achieved by means of binary splits, finding a decision criterion that best separated the multi-class data at each node into two groups. The decision criterion for this binary division was based on the attribute (e.g. feature) that returned the highest information gain. In addition, the classification of the data was based not only on one individual tree but on the weighted majority of multiple decision trees (Opitz & Maclin, 1999). With this bootstrap-aggregating approach, the effects of overfitting were reduced, improving generalization. The importance of each feature in the classification of the affective body movements was obtained from the most informed tree.

Four decision trees were considered for the classification of emotion using computed and behavioural ratings, which differed in the predictors used: (1) the postural and kinematic features averaged over time and *keypoints*; (2) the postural and kinematic features averaged over *keypoints* while keeping the temporal information; (3) nine behavioural ratings (i.e. all behavioural ratings excluding the emotional and action category ratings) and (4) only the six behavioural ratings that represented kinematic or postural aspects of the body movement (i.e. excluding the emotional and action category ratings as well as the ratings of emotional intensity,

valence and familiarity). These different sets of descriptors were used as input for each decision tree classifier, respectively.

In addition, two more decision trees were performed to investigate whether a given body part most significantly contributed to the distinction between emotions. Specifically, the first tree examined whether there were lateral asymmetries in the way the body expressed affect. This tree used as predictors the average value of the all the keypoint positions at the centre (i.e. nose and neck), at the left (i.e. left shoulder, elbow, wrist, hip, knee and ankle) and right side of the body (i.e. right shoulder, elbow, wrist, hip, knee and ankle), respectively, of each video. The second tree used fourteen descriptors representing the average keypoint location of the fourteen body joints, for all the videos.

One-way repeated-measures ANOVA

A one-way repeated-measures ANOVA was conducted for each feature to investigate possible differences between emotional categories. Each ANOVA was, therefore, constituted of a four-level factor Emotion (i.e. Anger, Happiness, Neutral and Fear) and used as input the feature's averaged values of each video. In the cases where sphericity was violated, Greenhouse-Geisser correction was applied.

Results

Computational features

Our first question was whether (dis)similarities in the kinematic and postural features of body movements would reflect the affective categorical structure. To investigate this question, several features were defined from 56 body movement video clips and pairwise comparisons between stimuli values were computed. **Figure 1** illustrates the results of this representational

similarity analysis at the kinematic/postural level. The resulting RDMs were arranged in the same order as the four-emotional categories RDM (see **Fig. 1**, upper left corner). The (dis)similarity structure of symmetry showed a clear dissociation between neutral and the rest of the affective body movements. Likewise, the neutral condition presented more similarities to itself with respect to shoulder ratio and limb angles than to the other emotional classes. For limb angles, the fearful condition also showed a high degree of within-category similarity and between-category dissimilarity. Kinematic RDMs such as the ones for velocity, acceleration and vertical movement did not reflect a clear differentiation between emotional categories.

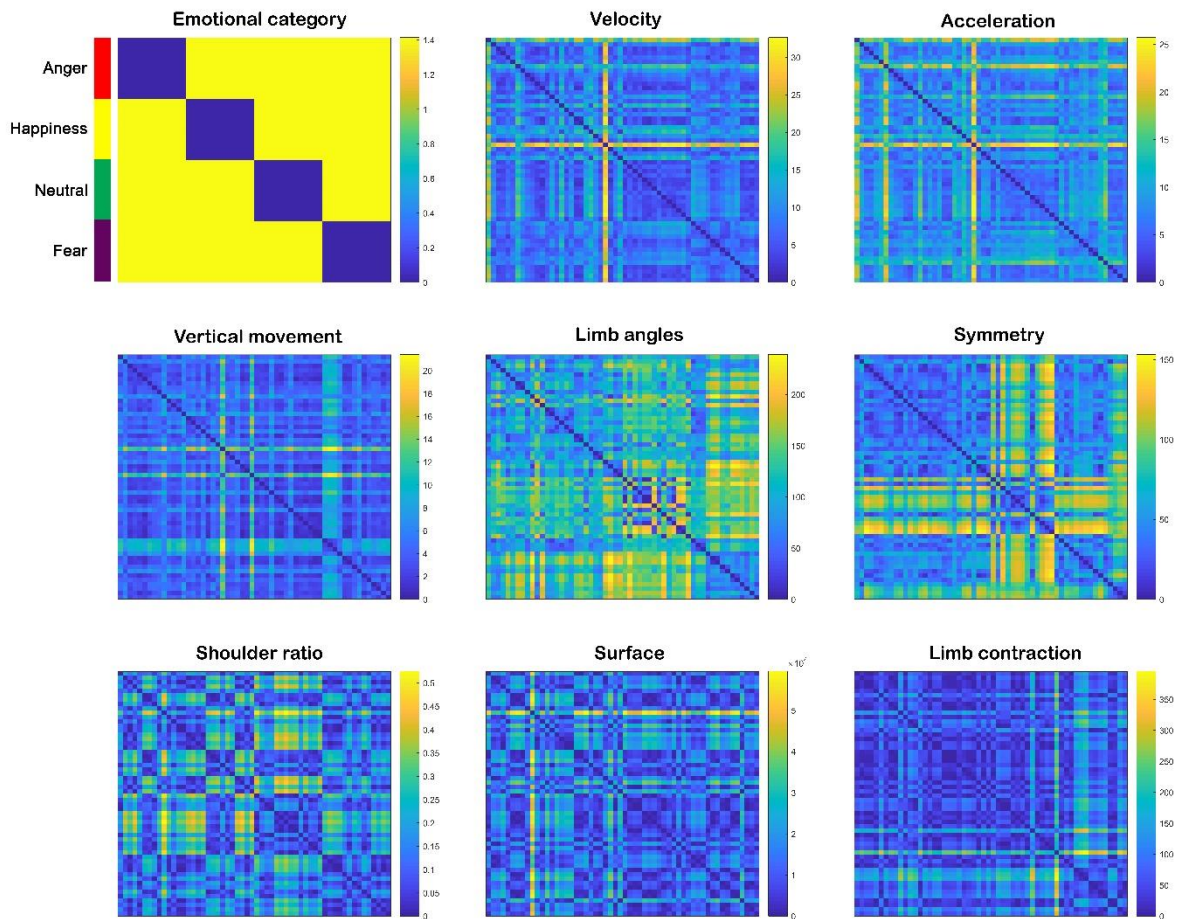


Figure 1. Representational dissimilarity matrices of the kinematic and postural features. The RDMs represent pairwise comparisons between 56 stimuli with regard to the

kinematic (i.e. velocity, acceleration and vertical movement) and postural (i.e. limb angles, symmetry, shoulder ratio, surface and limb contraction) computed features averaged over time (see Supplementary Materials for more information). The dissimilarity measure reflects Euclidean distance, with blue indicating strong similarity and yellow strong dissimilarity. Colour lines in the upper left corner indicate the organization of the RDMs with respect to the emotional category (anger: red; happiness: yellow; neutral: green; fear: purple) of the video stimuli.

To investigate whether the observed differences between emotional categories were significant, a one-way repeated-measures ANOVA was computed, separately, for each feature. Velocity showed a significant main effect of emotion ($F(1.764, 22.926) = 4.835$, $p = .021$, $\eta_p^2 = .271$), with the angry ($M = 3.4$, 95% CI = [2.54, 4.26]), happy ($M = 3.71$, 95% CI = [2.51, 4.91]) and fearful ($M = 3.14$, 95% CI = [2.61, 3.67]) conditions significantly faster than the neutral one ($M = 1.9$, 95% CI = [1.58, 2.21]) (see **Fig. SR1** in Supplementary Results). Acceleration also presented a main effect of emotion ($F(3, 39) = 5.202$, $p = .004$, $\eta_p^2 = .286$), with anger ($M = 4.13$, 95% CI = [3.35, 4.91]), happiness ($M = 4.11$, 95% CI = [3.16, 5.06]) and fear ($M = 4.01$, 95% CI = [3.42, 4.59]) showing significantly higher acceleration values than neutral movements ($M = 2.58$, 95% CI = [2.3, 2.87]). A significant main effect of emotion was also observed in the case of vertical movement ($F(3,39) = 3.226$, $p = .033$, $\eta_p^2 = .199$), where angry expressions ($M = .39$, 95% CI = [.15, .63]) presented more vertical displacement than fearful ones ($M = -.28$, 95% CI = [-.64, .09]). Limb angles showed a significant main effect of emotion ($F(3,39) = 13.499$, $p < .001$, $\eta_p^2 = .509$), with the limbs in the angry ($M = 141.87$, 95% CI = [135.61, 148.12]), happy ($M = 148.36$, 95% CI = [142.11, 154.6]) and neutral ($M = 140.39$, 95% CI = [135.76, 145.01]) conditions being significantly less flexed than in fear ($M = 121.27$, 95% CI = [112.32, 130.21]). The main effect of emotion was significant in the case of symmetry ($F(3,39) = 7.372$, $p < .001$, $\eta_p^2 = .362$), with angry ($M = 33.27$, 95% CI = [27.84, 38.7]) and happy ($M = 32.26$, 95% CI = [26.27, 38.26]) movements being significantly less symmetrical

than neutral ones ($M = 50.68$, 95% CI = [41.92, 59.43]). The significant main effect of emotion in shoulder ratio ($F(3,39) = 17.416$, $p < .001$, $\eta_p^2 = .573$) revealed that angry ($M = .49$, 95% CI = [.42, .57]), happy ($M = .5$, 95% CI = [.41, .59]) and fearful bodies ($M = .51$, 95% CI = [.46, .55]) were significantly more extended than neutral ones ($M = .69$, 95% CI = [.65, .74]). Surface presented a significant main effect of emotion ($F(3,39) = 3.712$, $p = .019$, $\eta_p^2 = .22$), with happy ($M = 47054.6$, 95% CI = [38966.71, 55142.48]) being significantly more extended than neutral movements ($M = 36090.46$, 95% CI = [32143.75, 40037.16]). Finally, the significant main effect of emotion for limb contraction ($F(3,39) = 10.410$, $p < .001$, $\eta_p^2 = .445$) revealed that angry ($M = 756.99$, 95% CI = [723.95, 790.02]), happy ($M = 751.89$, 95% CI = [733.97, 769.82]) and neutral bodies ($M = 803.6$, 95% CI = [763.79, 843.42]) were significantly more extended than the fearful ones ($M = 684.63$, 95% CI = [648.65, 720.61]).

To examine whether the kinematic and postural attributes relate to each other and/or to the emotional categories, pairwise comparisons were computed between the corresponding matrices (**Fig. 2**; see **Table SR1** in Supplementary Results for correlation and p-values; below only p-values corrected for multiple comparisons are reported). Interestingly, both postural and kinematic RDMs were positively correlated, although weakly, with the emotion RDM with the exception of vertical movement and surface: velocity ($r(1538) = .094$, $p = .002$), acceleration ($r(1538) = .101$, $p = .001$), limb angles ($r(1538) = .251$, $p < .001$), symmetry ($r(1538) = .262$, $p < .001$), shoulder ratio ($r(1538) = .185$, $p < .001$), and limb contraction ($r(1538) = .112$, $p < .001$). The feature matrices that most strongly correlated with the emotion RDM were, therefore, limb angles and symmetry. Kinematic RDMs overall correlated more strongly among each other, specially velocity and acceleration ($r(1538) = .768$, $p < .001$) while postural matrices showed weaker correlations. Among these, limb angles and symmetry ($r(1538) = .377$, $p < .001$), and shoulder ratio and surface ($r(1538) = .547$, $p < .001$) presented the relatively strongest

correlations. The relationship between postural and kinematic matrices was weak and often negative.

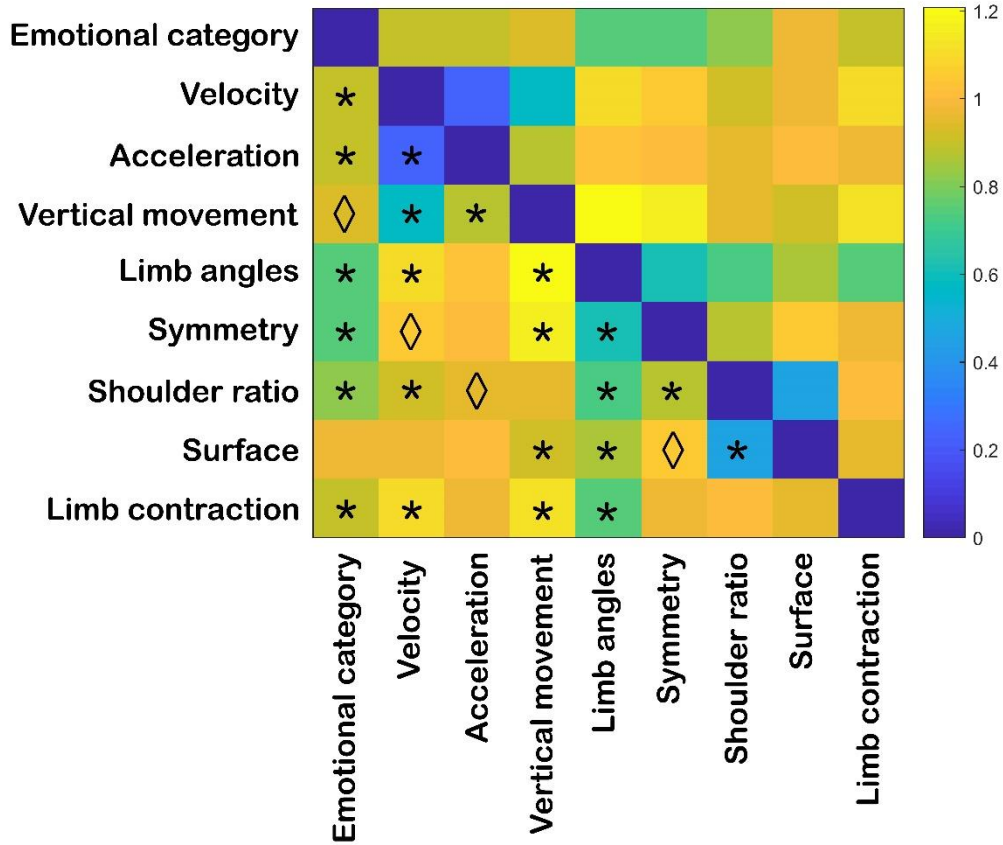


Figure 2. Correlation between representational dissimilarity matrices of kinematic and postural features. The RDM represents the level of (dis)similarity between each of the kinematic (i.e. velocity, acceleration and vertical movement) and postural (i.e. limb angles, symmetry, shoulder ratio, surface and limb contraction) matrices (see Fig. 1). Distances are indicated in 1-Spearman's correlation values, with blue indicating strong similarity and yellow strong dissimilarity. Asterisks and rhombi below the diagonal indicate significant correlations after Bonferroni correction and correlations that presented significant uncorrected p-values, respectively ($\alpha_{\text{bonf}} = .05/9$, with nine comparisons per feature; see Table SR1 in Supplementary Results for correlation and p-values).

To further investigate the relationship between affective states and kinematic and postural attributes of body movement, two decision tree classifiers were trained and tested. The

eight computed features were used as predictors and the emotional categories as the predicted class. The two models differed in whether or not the feature descriptors kept the temporal information of the movement. The model using features averaged over time provided a classification accuracy of 61% and showed that the angles between the limbs, symmetry, and the vertical displacement of the body joints are the most relevant features for the classification of emotion from body movements (see **Fig. 3**). When using feature descriptors that kept the temporal information (e.g. using information from individual frames), limb angles still appeared as the most relevant predictor, together with shoulder ratio and limb contraction. Importantly, this second model gave the higher accuracy of 84% (see **Fig. SR2** in Supplementary Results for an overview of predictor importance for these two models).

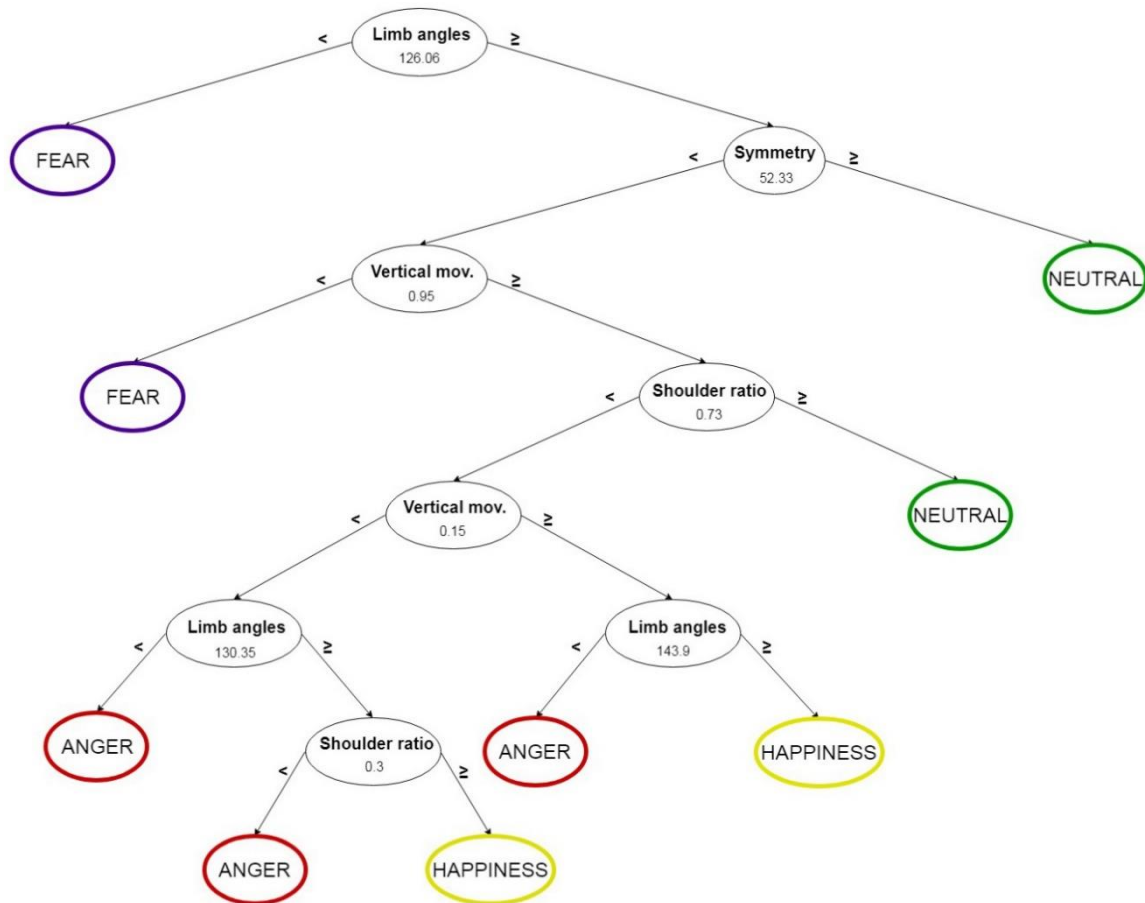


Figure 3. Emotional classification from postural and kinematic features averaged over time. A decision tree classifier was trained and tested (classification accuracy of 61%, with chance level at 25%) with the eight computed features as predictors and the four emotional categories as the predicted class. Kinematic features: velocity, acceleration and vertical movement. Postural features: limb angles, symmetry, shoulder ratio, surface and limb contraction.

Two more decision trees were performed that investigated whether a specific body part was most responsible for the recognition of emotion. This analysis revealed that the left side of the body was more relevant than the right side or the head/nose position in the distinction between emotions (classification accuracy > 90%, see **Fig. SR3.A** in Supplementary Results). A more detailed examination showed that the wrists, especially the left one, were the most

important body parts in the classification of affect (classification accuracy > 95%, see **Fig. SR3.B** in Supplementary Results).

Behavioural ratings

A further goal of this study was to investigate the (dis)similarity of different emotional body movements with regard to the perceived kinematic and postural features. For this purpose, 30 participants answered six questions concerning kinematic (i.e. amount of movement, fast movement, vertical movement, direction of the movement) and postural (i.e. body contraction, symmetry) aspects of the movement. To gain more insight on their perception of the stimuli, five more questions were asked about emotional- (i.e. emotional category, intensity, familiarity, valence) and action-related traits (i.e. action category) of the stimuli (see Methods and Supplementary Materials for more information on the behavioural task). **Figure 4** shows the average perceptual (dis)similarity scores across participants for each of the possible combinations of the 56 videos, for each rating, respectively. The (dis)similarity structure of kinematic-related ratings, such as amount of movement, fast and vertical movement, showed a relatively clear dissociation between neutral and the rest of the affective body movements. Likewise, the ratings for the neutral condition presented more within-category similarity with respect to contraction and symmetry than to the other emotional classes. This distinction was also clearly marked in the emotional intensity matrix, where the neutral category presented high within-category similarity and between-category dissimilarity. The four emotional categories did not show differences with regard to emotional intensity. However, participants reported having different degree of familiarity across categories: anger, neutral and fear presented within-category consistence while this was not the case for happiness. With respect to between-category comparisons for familiarity, only anger and fear showed some degree of similarity, whereas dissimilarity dominated the rest of the comparisons. A high within-category similarity was observed for valence. As could be expected, happiness displayed a strong dissimilarity to

fear and anger with respect to valence while the latter ones presented higher similarity. For the forward/away rating, the fearful condition also showed a high degree of within-category similarity and between-category dissimilarity while the other categories were more similar within and between each other.

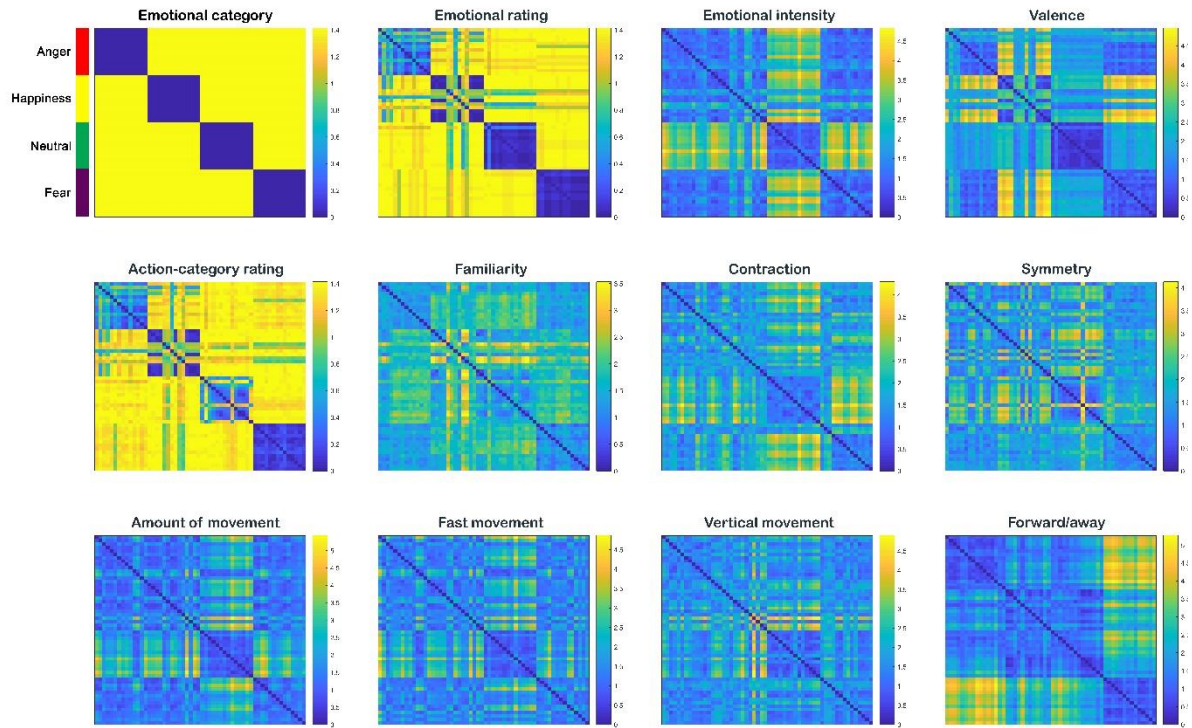


Figure 4. Representational dissimilarity matrices of the behavioural ratings. The RDMs represent pairwise comparisons between the 56 stimuli with regard to the each of the behavioural ratings (see Supplementary Materials for more information). The dissimilarity measure reflects Euclidean distance, with blue indicating strong similarity and yellow strong dissimilarity. Colour lines in the upper left corner indicate the organization of the RDMs with respect to the emotional category (anger: red; happiness: yellow; neutral: green; fear: purple) of the video stimuli.

To examine whether the kinematic-, postural-, emotional- and action-related ratings may correlate to each other and/or to the emotional categories, pairwise comparisons were computed between the corresponding matrices (see **Fig. 5** and **Table SR2** in Supplementary

Results for correlation and p-values; below only p-values corrected for multiple comparisons are reported). The behavioural rating on emotional categories correlated positively with all the behavioural ratings, showing the strongest correlations with valence ($r(1538) = .580, p < .001$) and action-category ratings ($r(1538) = .887, p < .001$). Participant's emotional ratings correlated positively with the emotional categories as well ($r(1538) = .728, p < .001$). Indeed, participants classified the affective body movements with high accuracy (see **Fig. SR4** in Supplementary Results for an inspection of the confusion matrix, also **Fig. 4** for the perceptual similarity RDM with respect to emotion). Neutral and fear were the most accurately recognised categories with 97 and 98% correct classification rate, respectively, while movements intended to express happiness had the lowest correct emotion attribution (78%), being most often confused with neutral body movements. In addition, the emotional categories correlated positively with all the remaining behavioural ratings, the strongest correlations being with action-category ratings ($r(1538) = .720, p < .001$) and forward/away ($r(1538) = .535, p < .001$). As with the computed features, kinematic ratings correlated more strongly among each other while postural ratings exhibited weaker correlations. The comparison between postural and kinematic ratings showed overall moderate to weak positive correlations.

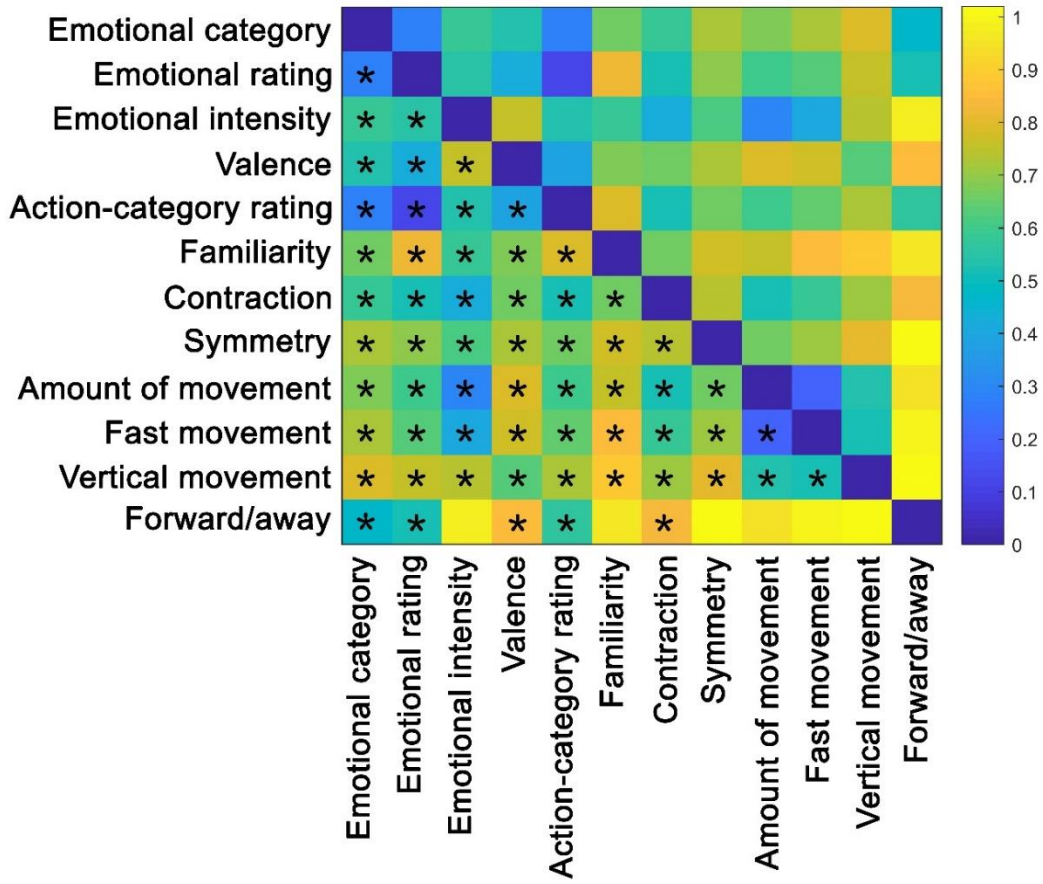


Figure 5. Correlation between representational dissimilarity matrices of the different behavioural ratings. The RDM represents the level of (dis)similarity between each of the behavioural-rating matrices (see Fig. 4). Distances are indicated in 1-Spearman's correlation values, with blue indicating strong similarity and yellow strong dissimilarity. Asterisks below the diagonal indicate significant correlations after Bonferroni correction ($\alpha_{\text{bonf}} = .05/12$, with 12 comparisons per behavioural rating; see Table SR2 in Supplementary Results for correlation and p-values).

To further investigate possible contributions of the perceived kinematic, postural and emotional attributes to the classification of emotion, two decision tree classifiers were trained and tested. The classification of emotion was performed using either (1) all behavioural ratings excluding emotional category rating and action category rating or (2) only the six behavioural ratings that represented kinematic or postural aspects of the movement (i.e. excluding emotional

category, action category, emotional intensity, valence, and familiarity ratings). The first model gave a classification accuracy of 78% and showed that the ratings of forward/away, valence and emotional intensity are the most relevant descriptors for the classification of emotional body movements. When using only the behavioural ratings that presented computed counterparts, forward-away was again the most relevant descriptor, followed by amount of movement and symmetry. This second model gave the lower accuracy of 71% (see **Fig. SR5** in Supplementary Results for an overview of the predictor importance for these two models).

Comparison between physical and perceived kinematic and postural body attributes

In addition, this study aimed at investigating whether perceptual (dis)similarities in kinematic and postural attributes across videos could be predicted by corresponding (dis)similarities at the computational level. To investigate this relationship, each of the perceptual RDMs was correlated to every computed feature RDM (**Fig. 6**; see **Table SR3** in Supplementary Results for correlation and p-values). With regard to the behavioural assessments (see previous section), the matrices showing the strongest positive correlation to emotion categories belonged to emotional ratings, action categories and forward/away. In the case of the computed features (see result section: postural and kinematic features), the strongest positive correlations to emotion categories were found for postural features such as symmetry, limb angles and shoulder ratio.

The behavioural task was designed such that the ratings would reflect common usage and still correspond to some of the computed features (see **Fig. 6**, where these correspondences are indicated as red squares). When evaluating the actual correlations between the behavioural and computed attributes at the RDM level, some of the behavioural ratings were indeed correlated the highest to their computational counterpart, although not in all cases. The behavioural rating of symmetry indeed showed the highest correlation with the computed

symmetry. However, the behavioural rating of contraction (which conceptually is related to the computation features shoulder ratio, surface and limb contraction) correlated highest with symmetry. The ratings of amount of movement, fast movement and vertical movement, which are conceptually related to computed velocity, acceleration and vertical movement, respectively, were all correlated significantly between each other. However, the behavioural rating of amount of movement was also highly correlated with symmetry, whereas fast movement correlated the strongest with velocity. Also, the rating of vertical movement was actually more strongly correlated to velocity than its computed counterpart.

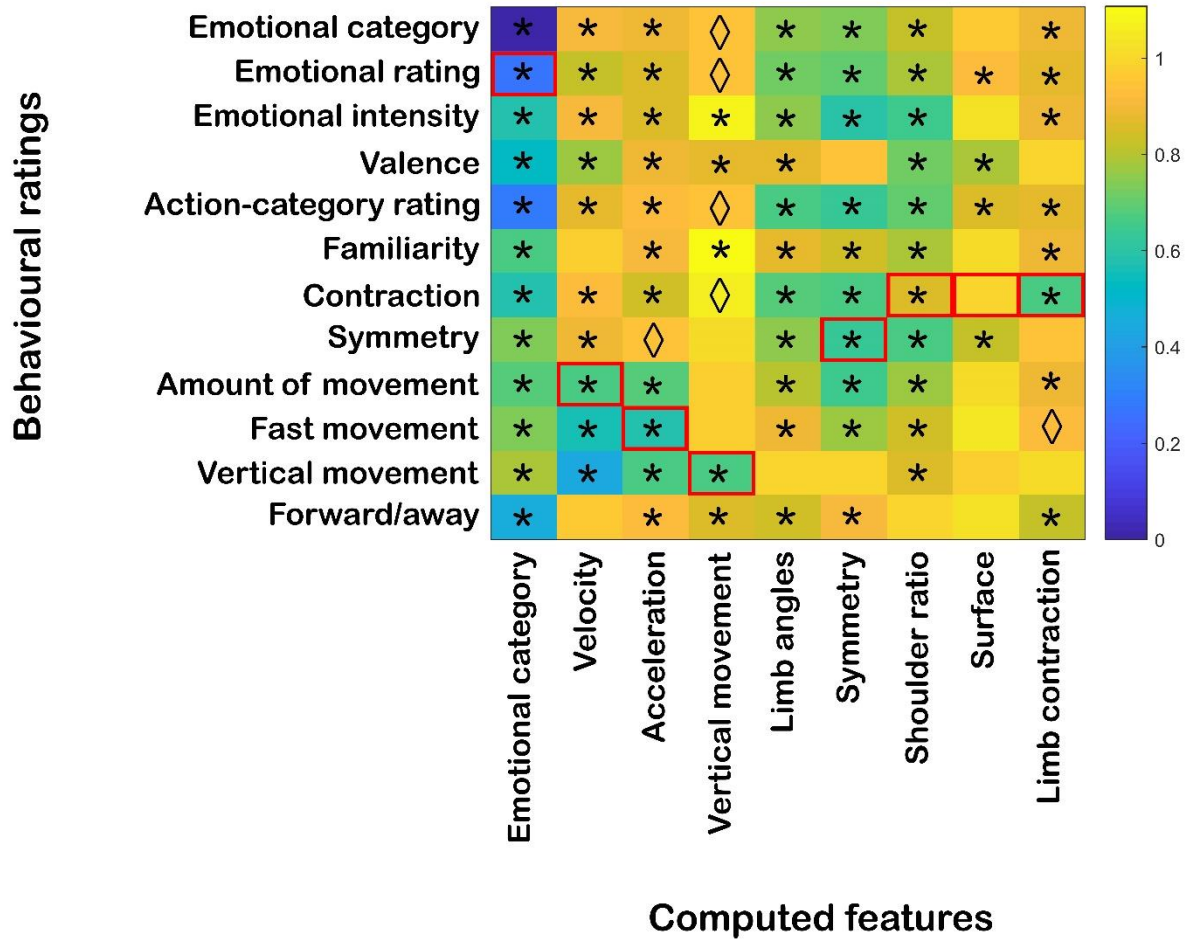


Figure 6. Average Spearman’s rank correlation between the behavioural-rating RDMs and the RDMs of the kinematic and postural features. Distances are indicated in 1-Spearman’s correlation values, with blue indicating strong similarity and yellow strong dissimilarity between behavioural and computed features. Asterisks and rhombi below the diagonal indicate significant correlations after Bonferroni correction and correlations that presented significant uncorrected p-values, respectively ($\alpha_{\text{bonf}} = .05/12$, with 12 comparisons per behavioural rating; see Table SR3 in Supplementary Results). Red boxes indicate the correspondence between computed features and behavioural ratings.

Discussion

This study used an innovative approach to investigate the role of kinematic and postural information in whole-body movement perception. Quantitative features of posture and movement derived from the position of the actors’ main joints were related to emotion

categories as well as to behavioural ratings of feature descriptors. Overall, postural rather than kinematic features seemed to discriminate better between different emotional body movements, both for the computed features as well as for the behavioural ratings. Among the postural descriptors, limb angles and symmetry appeared to be the most important cues. Moreover, adding time-related information to the computed features significantly improved the classification accuracy as well as changed the contribution of specific features in emotion classification. Finally, the degree of approach/avoidance behaviour seemed to be especially critical for the recognition of fear.

Computational features

Our first result from the RSA indicates that, overall, postural and not kinematic computational features differentiate between emotional categories (see **Fig. 1**). This finding was further supported by the correlation analyses between emotion and feature RDMs and the decision tree classifiers, where the postural features of limb angles and symmetry correlated best with emotional category (see **Fig. 2**) and were the two most important features for the classification tree (see **Fig. 3**). These results may indeed suggest that postural descriptors rather than kinematic ones play a more important role in the mechanism of decoding emotion from body movements.

Interestingly, the relevance of postural descriptors over kinematic ones was not only observed when removing the time information of the movement (i.e. averaging feature values over time) but even when preserving this information (see **Fig. SR2** in Supplementary Results). This suggests that time information may not be strictly necessary to distinguish between emotions, but provides additional information that can be used to solve difficult cases. Yet, the way in which these features were computed seems to be critical in determining the relevance of some attributes over others. Without the time information, limb angles, symmetry and the

vertical displacement of the joints were the most relevant features for emotion classification. Keeping the time information showed that limb angles still appeared as the most relevant predictor, but together with shoulder ratio and limb contraction. What is more important is that the classification accuracy was higher for the model whose feature descriptors preserved the time information, although using time-averaged features still provided above-chance classification performance.

Previous studies have also shown that the attribution of a specific emotion is indicated by the different contributions of postural and kinematic features (De Meijer, 1989; Wallbott, 1998). For example, it has been revealed that postural cues have a stronger influence compared to dynamic cues in anger (Aronoff, Woike, & Hyman, 1992) and fear expressions (Atkinson et al., 2007; Dittrich, Troscianko, Lea, & Morgan, 1996) than for happiness or sadness. In line with this finding, our study found that limb angles and limb contraction are relevant in differentiating fear from other expressive movements (see **Fig. SR1** in Supplementary Results). The amount of vertical displacement seems, however, important in discerning fear from anger. Interestingly, velocity, acceleration and shoulder ratio play a role in differentiating emotional from non-emotional body movements. These results may not be apparent from the representational similarity analyses in the case of velocity and acceleration. A possible explanation could be that while the former used time- and keypoint-averaged data, the latter only time-averaged information.

Behavioural features

Our second result concerns the behavioural features and how they relate to emotion categories. This provides a picture of which features may predominantly guide subjective emotion recognition, yet not providing a direct link. Here, the examination of participants' perceptual judgements about kinematic as well as contraction and symmetry descriptors revealed a clear

discrimination between neutral and the rest of the affective categories (see **Fig. 4**). Interestingly, a strong dissimilarity was observed between fear and other emotions with regard to approach/avoidance behaviour. This may explain the high recognition rate observed for fearful expressions, in line with previous literature showing that retreating behaviour is a diagnostic feature of fear (Atkinson et al., 2007). The importance of this attribute was further confirmed by the classification analysis, showing that this feature was the most relevant for the classification of emotion from perceptual ratings, together with contraction and symmetry (see **Fig. SR5** in Supplementary Results).

Relation between computational and behavioural features

As mentioned in previous sections, the distinction between different affective movements using computed kinematic features was unclear. However, participants' ratings on the corresponding attributes revealed a relatively clearer distinction between emotional and non-emotional movements (see **Fig. 4**). This gap between computed and perceptual descriptors was also reflected in the weak correlations between their respective matrices (see **Fig. 6**). It could be that the approach followed to calculate the kinematic features does not characterize well how people process movement information for the distinction between different affective movements.

With regard to postural cues, more similarities were found between the computed and perceptual features. The clearest example was observed with symmetry. In the case of the behavioural rating of “contraction”, three different computed counterparts were defined: shoulder ratio, surface and limb contraction. However, the behavioural rating of this attribute was closer to the representation of limb contraction and limb angles than to surface or shoulder ratio. As this example shows, a specific attribute can be computed in indeed multiple ways.

High dimensionality

An important difficulty in studying body movement and posture is the fact that the body is a complex high-dimensional stimulus (Roether et al., 2009). Some studies have approached this issue by selecting a body part that most significantly contributed to the distinction between emotions (Glowinski et al., 2015; Pollick et al., 2001; Roether, Omlor, & Giese, 2008; Wallbott, 1998). This study is in agreement with previous findings showing that left side of the body is more emotionally expressive than the right (Roether et al., 2008), specially the wrist (see **Fig. SR3** in Supplementary Results). Although the selection of a body part would have reduced the dimensionality of the defined features, the aim of the current study was to investigate how whole-body movements, rather than its parts, convey emotion.

Taking a more methodological and systematic approach, other authors investigating affective movements, but also intention and motor control, have used data-reduction methods such as PCA (Santello, Flanders, & Soechting, 2002; Troje, 2002; Yacoob & Black, 1999), factor analysis (Ivanenko, Cappellini, Dominici, Poppele, & Lacquaniti, 2005; Ivanenko, Poppele, & Lacquaniti, 2004; Soechting, 1997; Tresch, Cheung, & d'Avella, 2006) or blind-source separation algorithm (Roether et al., 2009). In this study, several postural and kinematic features were selected for their demonstrated importance in previous literature. This approach was followed since relevant body descriptors obtained from perceptual experiments have been shown to be in good agreement with those extracted from data-driven approaches (Roether et al., 2009).

Conclusion

This study made a first attempt to relate features of body posture and movement to emotional categories as well as to observers' subjective perception of feature descriptors. By the recent advances in machine learning algorithms for person detection, it is now possible to

quantitatively derive core feature of human movement and investigate their role in human perception. We showed that postural rather than kinematic features seem to discriminate better between different emotional body movements, both for the computed features as well as for the behavioural ratings. A next step is the further understanding of these features in relation to the different brain regions involved in body perception. Insights in these mechanisms will have a crucial impact on our understanding of affect and social interaction, but also in many areas of society, including law enforcement and security, games and entertainment, education, the arts (Camurri et al., 2016) but, most importantly, health care (Kleinsmith & Bianchi-Berthouze, 2012). Patients suffering from disorders of affective communication, such as autism and schizophrenia, will directly benefit from the application of this knowledge to rehabilitation programmes focused on emotional recognition and normal social functioning.

Acknowledgments

This work was supported by the European Research Council (ERC) FP7-IDEAS-ERC (Grant agreement number 295673; Emobodies), by the Future and Emerging Technologies (FET) Proactive Programme H2020-EU.1.2.2 (Grant agreement 824160; EnTimeMent) and by the Industrial Leadership Programme H2020-EU.1.2.2 (Grant agreement 825079; MindSpaces). We would like to thank Caoimhe Moran for the assistance with behavioural data acquisition.

Author Contributions

All authors contributed to the study concept and design. M. P. S. performed the data analysis under the supervision of M.V. M. P. S. performed the interpretation of the results and drafted

the manuscript, and M. V. and B. d. G. provided critical feedback. All authors approved the final version of the manuscript for submission.

Competing interests

The authors declare no competing interests.

References

- Aronoff, J., Woike, B. A., & Hyman, L. M. (1992). Which are the stimuli in facial displays of anger and happiness? Configurational bases of emotion recognition. *J. Pers. Soc. Psychol.*, 62(6), 1050 - 1066.
- Atkinson, A. P., Tunstall, M. L., & Dittrich, W. H. (2007). Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition*, 104(1), 59-72.
- Camurri, A., Volpe, G., Piana, S., Mancini, M., Niewiadomski, R., Ferrari, N., & Canepa, C. (2016). The dancer in the eye: towards a multi-layered computational framework of qualities in movement. *Proceedings of the 3rd International Symposium on Movement and Computing*, 6.
- Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7291-7299.
- Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *J. Nonverbal Behav.*, 28(2), 117-139.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nat. Rev. Neurosci.*, 7(3), 242-249.
- de Gelder, B., Van den Stock, J., Meeren, H. K., Sinke, C. B., Kret, M. E., & Tamietto, M. (2010). Standing up for the body. Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neurosci. Biobehav. Rev.*, 34(4), 513-527.
- De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *J. Nonverbal Behav.*, 13(4), 247-268.
- Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25(6), 727-738.
- Glowinski, D., Mortillaro, M., Scherer, K., Dael, N., Volpe, G., & Camurri, A. (2015). Towards a minimal representation of affective gestures. *International Conference on Affective Computing and Intelligent Interaction*, 498-504.
- Ivanenko, Y., Cappellini, G., Dominici, N., Poppele, R., & Lacquaniti, F. (2005). Coordination of locomotion with voluntary movements in humans. *J. Neurosci.*, 25(31), 7238-7253.

- Ivanenko, Y., Poppele, R., & Lacquaniti, F. (2004). Five basic muscle activation patterns account for muscle activity during human locomotion. *J. Physiol.*, 556(1), 267-282.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Percept. Psycho.*, 14(2), 201-211.
- Kleinsmith, A., & Bianchi-Berthouze, N. (2012). Affective body expression perception and recognition: A survey. *IEEE Trans. Affect. Comput.*, 4(1), 15-33.
- Kret, M., Pichon, S., Grèzes, J., & de Gelder, B. (2011). Similarities and differences in perceiving threat from dynamic faces and bodies. An fMRI study. *NeuroImage*, 54(2), 1755-1762.
- Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., . . . Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, 60(6), 1126-1141.
- Loh, W.-Y. (2002). Regression tress with unbiased variable selection and interaction detection. *Stat. Sin.*, 361-386.
- Niewiadomski, R., Mancini, M., Piana, S., Alborno, P., Volpe, G., & Camurri, A. (2017). Low-intrusive recognition of expressive movement qualities. *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 230-237.
- Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Comput. Biol.*, 10(4), e1003553.
- Opitz, D., & Maclin, R. (1999). Popular ensemble methods: An empirical study. *Journal of artificial intelligence research*, 11, 169-198.
- Paterson, H. M., Pollick, F. E., & Sanford, A. J. (2001). The role of velocity in affect discrimination. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 23(23).
- Patwardhan, A. (2017). Three-Dimensional, Kinematic, Human Behavioral Pattern-Based Features for Multimodal Emotion Recognition. *Multimodal technol. interact.*, 1(3), 19.
- Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, 162(1-2), 8-13.
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in neuroinformatics*, 2, 10.
- Piana, S., Stagliano, A., Odone, F., Verri, A., & Camurri, A. (2014). Real-time automatic emotion recognition from body gestures. *arXiv preprint arXiv:1402.5047*.
- Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, 82(2), B51-B61.
- Roether, C. L., Omlor, L., Christensen, A., & Giese, M. A. (2009). Critical features for the perception of emotion from gait. *J. Vis.*, 9(6), 15-15.
- Roether, C. L., Omlor, L., & Giese, M. A. (2008). Lateral asymmetry of bodily emotion expression. *Curr. Biol.*, 18(8), R329-R330.
- Santello, M., Flanders, M., & Soechting, J. F. (2002). Patterns of hand motion during grasping and the influence of sensory guidance. *J. Neurosci.*, 22(4), 1426-1435.
- Sawada, M., Suda, K., & Ishii, M. (2003). Expression of emotions in dance: Relation between arm movement characteristics and emotion. *Percept. Mot. Skills*, 97(3), 697-708.
- Schindler, K., Van Gool, L., & de Gelder, B. (2008). Recognizing emotions expressed by body pose: A biologically inspired neural model. *Neural Netw.*, 21(9), 1238-1246.
- Soechting, M. S. J. F. (1997). Matching object size by controlling finger span and hand shape. *Somatosens. Mot. Res.*, 14(3), 203-212.
- Tresch, M. C., Cheung, V. C. K., & d'Avella, A. (2006). Matrix Factorization Algorithms for the Identification of Muscle Synergies: Evaluation on Simulated and Experimental Data Sets. *J. Neurophysi.*, 95(4), 2199-2212.

- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *J. Vis.*, 2(5), 2-2.
- Vaessen, M. J., Abassi, E., Mancini, M., Camurri, A., & de Gelder, B. (2018). Computational feature analysis of body movements reveals hierarchical brain organization. *Cerebr. Cortex*, 1, 10.
- Wallbott, H. G. (1998). Bodily expression of emotion. *Eur. J. Soc. Psychol.*, 28(6), 879-896.
- Yacoob, Y., & Black, M. J. (1999). Parameterized modeling and recognition of activities. *Comput. Vis. Image Und.*, 73(2), 232-247.

Supplementary Information for:

**Features and feelings. Objective and subjective features critical for
emotional recognition in whole-body movements**

Marta Poyo Solanas¹, Maarten Vaessen¹, & Beatrice de Gelder^{1,2,*}

¹ Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Limburg 6200 MD, The Netherlands.

² Department of Computer Science, University College London, London WC1E 6BT, UK.

* Address correspondence to Beatrice de Gelder, Department of Cognitive Neuroscience, Brain and Emotion Laboratory, Faculty of Psychology and Neuroscience, Maastricht University, Oxfordlaan 55, 6229 EV Maastricht, The Netherlands. Tel. +31 433881437. E-mail: b.degelder@maastrichtuniversity.nl

Supplementary Materials

Feature definition

To investigate the contribution of kinematic and postural information to the processing and recognition of emotional body movements, several quantitative features were computed giving their importance in previous work (for a review see Kleinsmith & Bianchi-Berthouze, 2012). These features include velocity, acceleration, vertical movement, symmetry, limb angles and three different computations of body contraction. However, there are different ways in which these features can be calculated. Here, a stimulus dataset comprising 56 affective body-movement videos expressing either anger, fear, happiness or a non-emotional expression were used. The x- and y-coordinates of a total of 18 *keypoints* that corresponded to the actor's main body joints were acquired using OpenPose (v1.0.1) (Cao, Simon, Wei, & Sheikh, 2017), for each of the 25 frames that constituted each video (see **Fig. SM1**). Due to the blurring of the face in our video clips, the estimation of the location of eyes and ears was often inaccurate. These *keypoints* were disregarded for further analyses, although the *keypoint* corresponding to the nose was kept as a reference for the position of the head. Velocity was calculated as the amount of displacement of each *keypoint* for adjacent frames. Acceleration was derived from the difference in the amount of movement of each *keypoint* for adjacent frames. Vertical movement referred to the amount of displacement of each *keypoint* in the y-axes between adjacent frames. The feature defined as “limb angles” was calculated as the angle between two adjacent body segments, including the angles for the elbows, knees, shoulders and hips. Symmetry was computed as the difference in position of each pair of joints (i.e. one on the left side, the other on the right) with respect to the axis that divides the body vertically by the nose. Shoulder ratio was defined as the amount of extension of the body joints with respect to the shoulders. Surface was computed as the multiplication between the total body extension in the

x-axis and the extension in the y-axis. Limb contraction referred to the feature where the distances of the limbs (wrists and ankles) to the head were estimated. Although each feature was calculated within each frame, the results presented in the main text correspond to features whose time information was averaged. In the supplementary results, the dissimilarity matrices of the features preserving the time-related information are shown (see **Fig. SR6**). Finally, emotional categories simply denoted whether a video belonged to a category (anger, happiness, neutral or fear) or not.



Figure SM1. Example of OpenPose skeleton on a video from our data set.

Experimental questions for the behavioural part of the experiment

One of the goals of this study was to investigate the (dis)similarity of different emotional body movements with regard to the perceived kinematic and postural attributes. For this purpose, 30 participants answered six questions concerning kinematic (i.e. amount of movement, fast movement, vertical movement, direction of the movement) and postural (i.e. body contraction, symmetry) aspects of the movement. To gain more insight on their perception of the stimuli, five more questions were asked about emotional- (i.e. emotional category, intensity, familiarity, valence) and action- related (i.e. action category) traits of the stimuli. The movement- and postural- related questions were rated on a seven-point scale, whereas the emotional and action categorization ones were forced-choice. The order and the content of the questions can be seen in **Table SM1**.

Table SM1. Experimental questions and ratings for the behavioural part of the experiment

Questions	Answer
1.- How much body movement is there in the video?	Little movement = 1; A lot of movement = 7
2.- How fast is the body movement in the video?	Very slow = 1; Very fast = 7
3.- How much vertical movement is shown in the video?	Little movement = 1; A lot of movement = 7
4.- How much body contraction is there in the body movement?	Little contraction = 1; A lot of contraction = 7
5.- How symmetrical is the body movement?	Little symmetrical = 1; Very symmetrical = 7
6.- Is the movement directed towards you or away from you?	Away from you = 1; Towards you = 7
7.- Which action do you think is represented by the movement?	Walking, threatening, celebrating, coughing, self-protecting
8.- How familiar is the body movement to you?	Very unfamiliar = 1; Very familiar = 7
9.- What is the valence of the body movement?	Very negative = 1; Very positive = 7
10.- Which emotion do you think the person is expressing?	Anger, happiness, neutral, fear
11.- How intense is the emotion being expressed?	Little intense = 1; Very intense = 7

Supplementary Results

Emotional recognition and validity of the stimuli

Participants accurately classified the emotion expressed by the body movements, indicating the validity of the stimuli for the purpose of the study. Fear and neutral conditions were the best recognized, whereas happiness was the worst, being most often confused with neutral body expressions (see **Fig. SR4** in Supplementary Results). The observation that emotional intensity ratings were similar across emotions excludes this factor as a possible confound in the misclassification of happiness (see **Fig. 4** in main text). In addition, the movements conveying this emotion had similar kinematic and postural within-category similarity (see **Fig. 1** in main text). A possible explanation for this selective misclassification could be the different level of familiarity that participants presented with each affective movement. An inspection of the matrix for familiarity ratings (**Fig. 4** in main text) revealed that while fear, neutral and anger categories had high within-category similarity, the happy condition displayed less within-category consistency. Another possible explanation could be the existence of individual differences in affective expression between actors. Indeed, a closer look at participants' emotional ratings indicated that the lower recognition accuracies were specifically attributed to two actors.

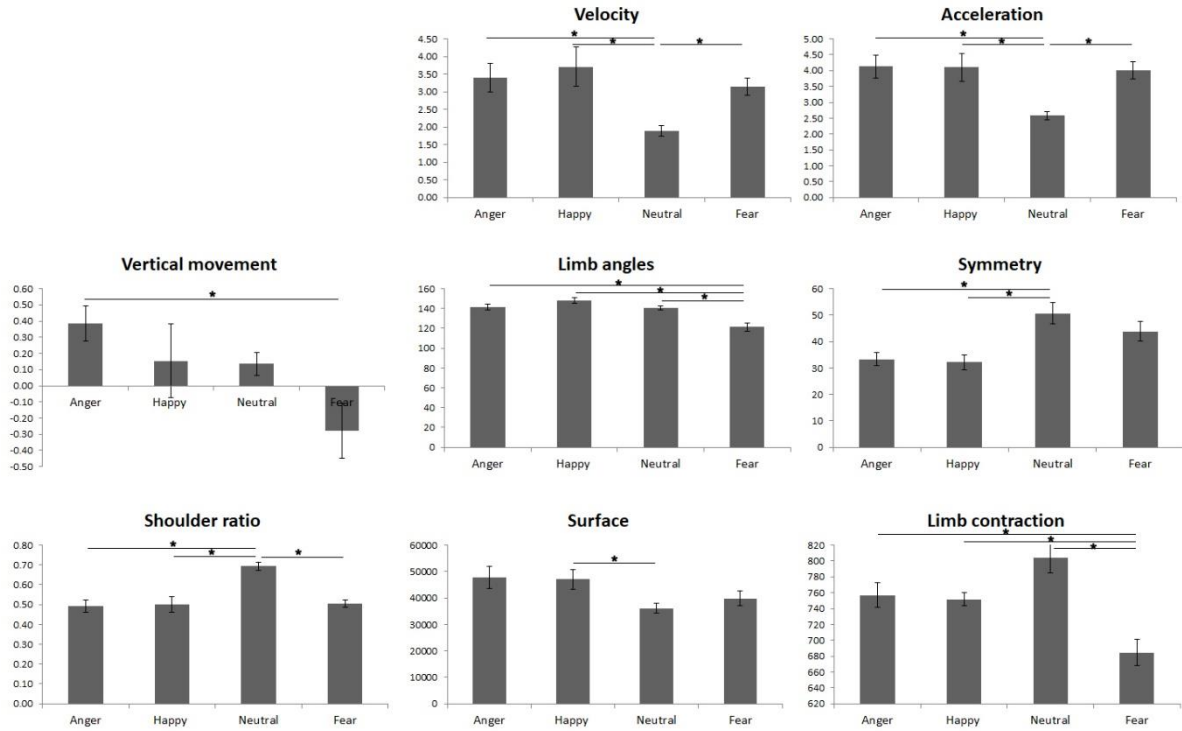


Figure SR1. Feature differences across emotions. For each feature, the averaged values for each video served as input for the ANOVA with a four-level factor Emotion (Anger, Happiness, Neutral and Fear). The lines and asterisks represent significant differences between emotions: **Velocity**: $F(1.764, 22.926) = 4.835$, $p = .021$, $\eta_p^2 = .271$, Anger vs. Neutral ($p = .010$), Happy vs. Neutral ($p = .049$), Fear vs. Neutral ($p = .002$); **Acceleration**: $F(3, 39) = 5.202$, $p = .004$, $\eta_p^2 = .286$, Anger vs. Neutral ($p = .005$), Happy vs. Neutral ($p = .047$), Fear vs. Neutral ($p = .002$); **Vertical movement**: $F(3,39) = 3.226$, $p = .033$, $\eta_p^2 = .199$, Anger vs. Fear ($p = .048$); **Limb angles**: $F(3,39) = 13.499$, $p < .001$, $\eta_p^2 = .509$, Anger vs. Fear ($p = .029$), Happy vs. Fear ($p < .001$), Neutral vs. Fear ($p = .002$); **Symmetry**: $F(3,39) = 7.372$, $p < .001$, $\eta_p^2 = .362$, Anger vs. Neutral ($p = .017$), Happy vs. Neutral ($p = .025$); **Shoulder ratio**: $F(3,39) = 17.416$, $p < .001$, $\eta_p^2 = .573$, Anger vs. Neutral ($p < .001$), Happy vs. Neutral ($p = .001$), Fear vs. Neutral ($p < .001$); **Surface**: $F(3,39) = 3.712$, $p = .19$, $\eta_p^2 = .22$, Happy vs. Neutral ($p = .29$); **Limb contraction**: $F(3,39) = 10.410$, $p < .001$, $\eta_p^2 = .445$, Anger vs. Fear ($p = .025$), Happy vs. Fear ($p = .003$), Neutral vs. Fear ($p = .002$).

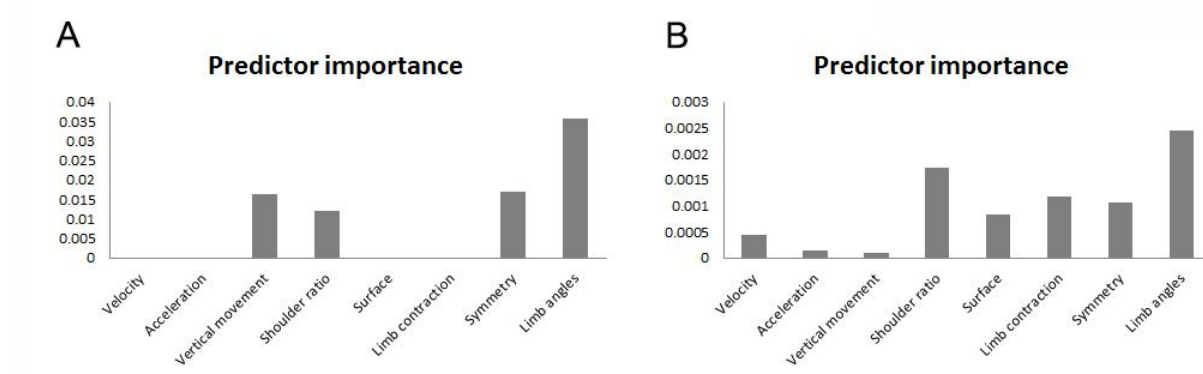


Figure SR2. Feature importance for the classification of emotion. **A)** Predictor relevance for the classification model where the postural and kinematic features were averaged over time; **B)** Predictor relevance for the classification model where the postural and kinematic features kept the temporal information.

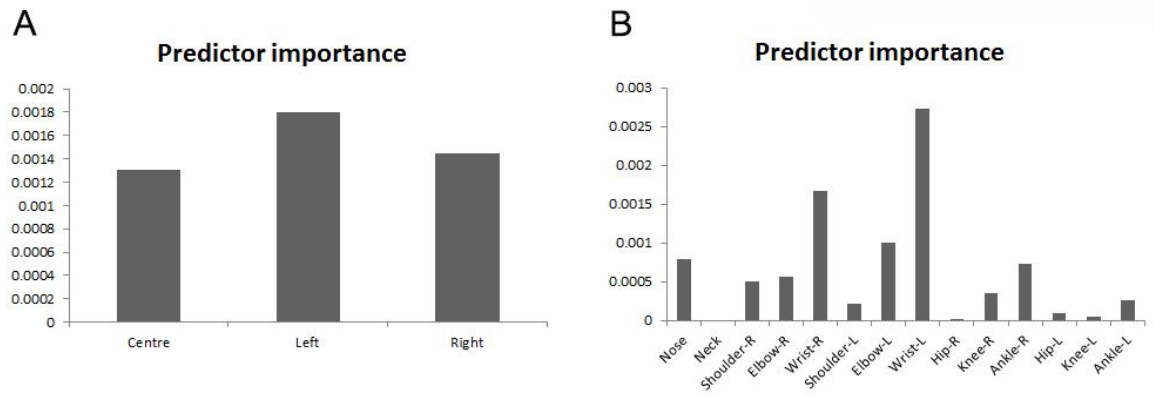


Figure SR3. Body part importance for the classification of emotion. **A)** Predictor relevance for the classification model using the average of the keypoints at the centre of the body (i.e. nose and neck), at the left (i.e. left shoulder, elbow, wrist, hip, knee and ankle) and right side (i.e. right shoulder, elbow, wrist, hip, knee and ankle); **B)** Predictor relevance for the classification model using all the fourteen body keypoints.

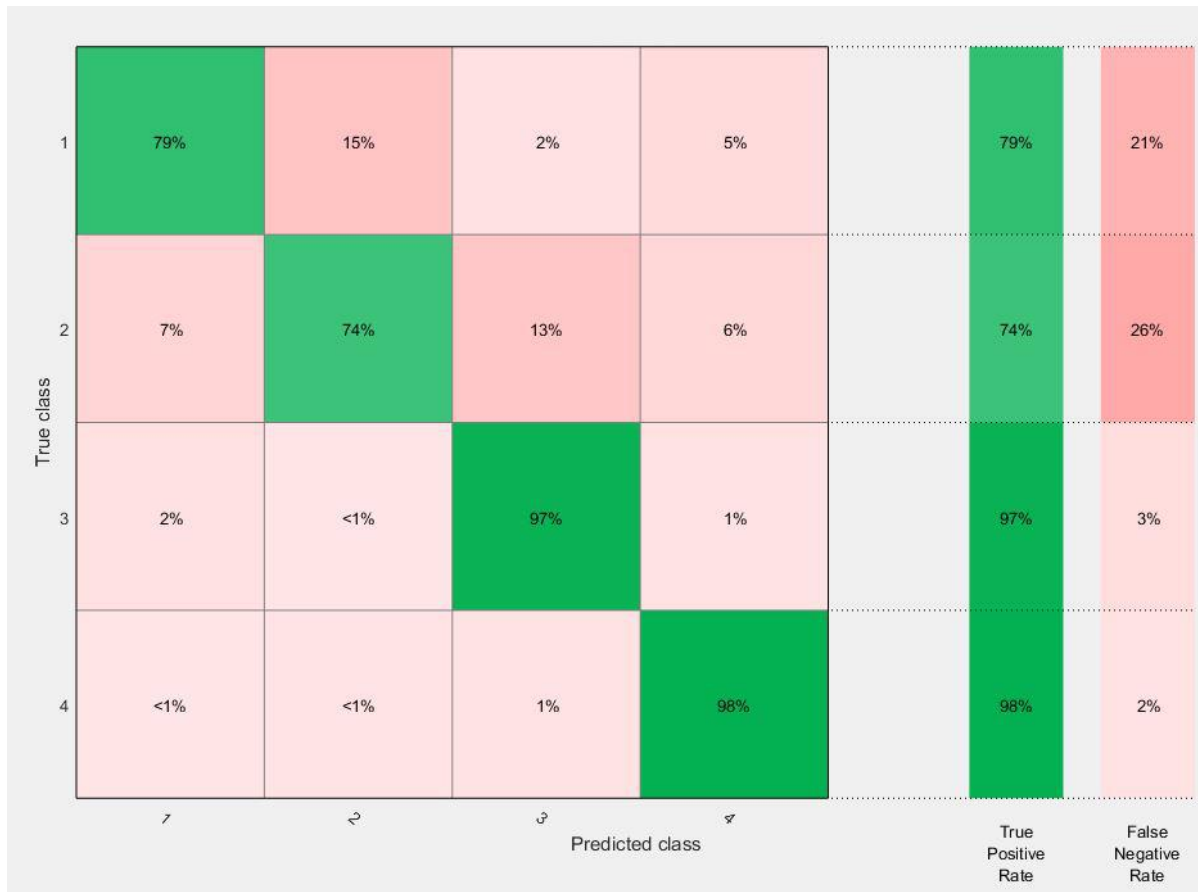


Figure SR4. Confusion matrix with participant’s emotional ratings as predictor variable (“True Class”) and the true emotional categories as outcome (“Predicted class”). The numbers in the y- and x- axes indicate the different emotional categories (1 = Anger; 2 = Happy; 3 = Neutral; 4 = Fear).

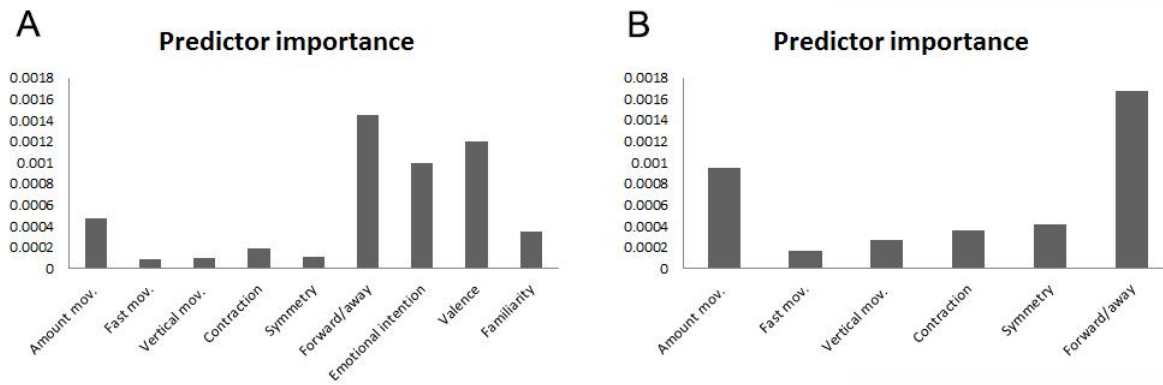


Figure SR5. Behavioural rating importance for the classification of emotion. **A)** Predictor relevance for the classification model where the ratings of postural, kinematic and emotional traits were included; **B)** Predictor relevance for the classification model where the only postural and kinematic ratings were included.

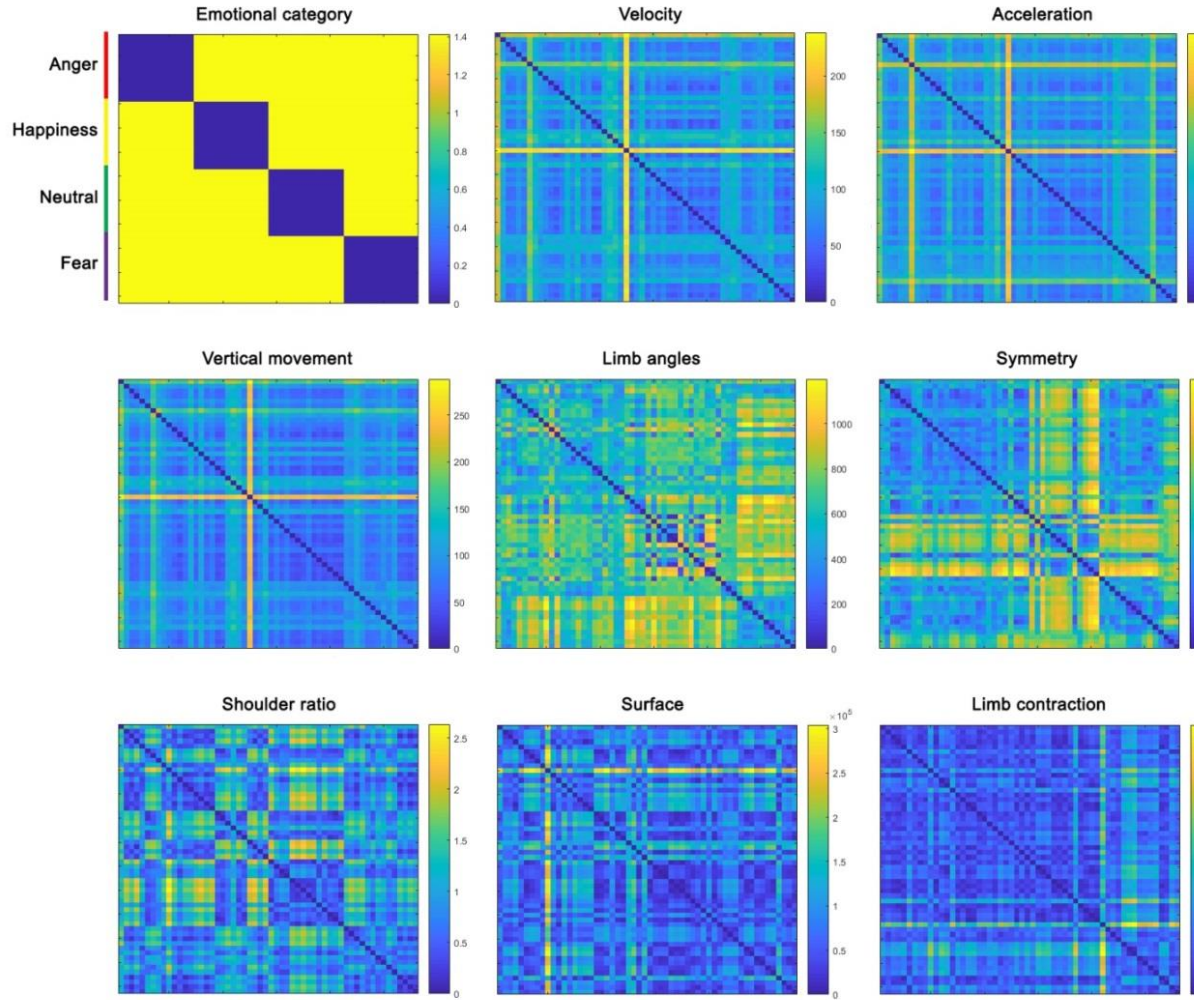


Figure SR6. Representational dissimilarity matrices of the kinematic and postural features preserving the time information. The RDMs represent pairwise comparisons between 56 stimuli with regard to the kinematic and postural computed features. The dissimilarity measure reflects Euclidean distance, with blue indicating strong similarity and yellow strong dissimilarity. Colour lines in the upper left corner indicate the organization of the RDMs with respect to the emotional category (anger: red; happiness: yellow; neutral: green; fear: purple) of the video stimuli.

TABLE SR1. Correlation between representational dissimilarity matrices of kinematic and postural features. The correlations were computed using Spearman's rank correlation. Both uncorrected and Bonferroni-corrected p-values are displayed for each comparison ($\alpha_{\text{bonf}} = 0.05/9$, with nine comparisons per feature).

	Emotional category			Velocity			Acceleration			Vertical movement		
	r	p-value	correct p-value	r	p-value	correct p-value	r	p-value	correct p-value	r	p-value	correct p-value
Velocity	.094	.000	.002									
Acceleration	.101	.000	.001	.768	.000	.000						
Vertical movement	.057	.025	.227	.417	.000	.000	.120	.000	.000			
Limb angles	.251	.000	.000	-.096	.000	.002	-.032	.210	.999	-.208	.000	.000
Symmetry	.262	.000	.000	-.054	.033	.294	-.009	.721	.999	-.155	.000	.000
Shoulder ratio	.185	.000	.000	.094	.000	.002	.051	.045	.404	.041	.111	.998
Surface	.036	.161	.999	.026	.305	.999	-.016	.543	.999	.087	.001	.006
Limb contraction	.112	.000	.000	-.104	.000	.000	.033	.200	.999	-.115	.000	.000

TABLE SR1 (continuation). Correlation between representational dissimilarity matrices of kinematic and postural features. The correlations were computed using Spearman's rank correlation. Both uncorrected and Bonferroni-corrected p-values are displayed for each comparison ($\alpha_{\text{bonf}} = 0.05/9$, with nine comparisons per feature).

	Limb angles			Symmetry			Shoulder ratio			Surface		
	r	p-value	correct p-value	r	p-value	correct p-value	r	p-value	correct p-value	r	p-value	correct p-value
Symmetry	.377	.000	.000									
Shoulder ratio	.274	.000	.000	.119	.000	.000						
Surface	.136	.000	.000	-.057	.026	.232	.547	.000	.000			
Limb contraction	.261	.000	.000	.029	.257	.999	-.019	.452	.999	.045	.077	.691

Table SR2. Correlation between representational dissimilarity matrices of the behavioural ratings. The correlations were computed using Spearman's rank correlation. Both uncorrected and Bonferroni-corrected p-values are displayed for each comparison ($\alpha_{\text{bonf}}=0.05/12$, with 12 comparisons per behavioural rating).

	Emotional Categories			Emotional Rating			Emotional Intensity			Valence		
	r	p-value	Corr. p-value	r	p-value	Corr. p-value	r	p-value	Corr. p-value	r	p-value	Corr. p-value
Emotional rating	.728	.000	.000									
Emotional intensity	.414	.000	.000	.455	.000	.000						
Valence	.466	.000	.000	.580	.000	.000	.246	.000	.000			
Action rating	.720	.000	.000	.887	.000	.000	.466	.000	.000	.602	.000	.000
Familiarity	.337	.000	.000	.181	.000	.000	.415	.000	.000	.323	.000	.000
Contraction	.414	.000	.000	.474	.000	.000	.576	.000	.000	.343	.000	.000
Symmetry	.268	.000	.000	.311	.000	.000	.389	.000	.000	.279	.000	.000
Amount of movement	.323	.000	.000	.394	.000	.000	.712	.000	.000	.207	.000	.000
Fast movement	.270	.000	.000	.370	.000	.000	.599	.000	.000	.229	.000	.000
Vertical movement	.213	.000	.000	.251	.000	.000	.257	.000	.000	.363	.000	.000
Forward-away	.535	.000	.000	.477	.000	.000	.018	.468	.999	.147	.000	.000

Abbreviation: Corr.: Bonferroni-corrected p-value

Table SR2 (continuation). Correlation between representational dissimilarity matrices of the behavioural ratings. The correlations were computed using Spearman's rank correlation. Both uncorrected and Bonferroni-corrected p-values are displayed for each comparison ($\alpha_{\text{bonf}} = 0.05/12$, with 12 comparisons per behavioural rating).

	Action Rating			Familiarity			Contraction			Symmetry		
	r	p-value	Corr. p-value	r	p-value	Corr. p-value	r	p-value	Corr. p-value	r	p-value	Corr. p-value
Familiarity	.216	.000	.000									
Contraction	.477	.000	.000	.345	.000	.000						
Symmetry	.340	.000	.000	.234	.000	.000	.252	.000	.000			
Amount of movement	.396	.000	.000	.245	.000	.000	.484	.000	.000	.337	.000	.000
Fast movement	.347	.000	.000	.155	.000	.000	.425	.000	.000	.288	.000	.000
Vertical movement	.269	.000	.000	.121	.000	.000	.289	.000	.000	.188	.000	.000
Forward-away	.441	.000	.000	.031	.219	.999	.158	.000	.000	-.020	.428	.999

Abbreviation: Corr.: Bonferroni-corrected p-value

Table SR2 (continuation). Correlation between representational dissimilarity matrices of the behavioural ratings. The correlations were computed using Spearman's rank correlation. Both uncorrected and Bonferroni-corrected p-values are displayed for each comparison ($\alpha_{\text{bonf}} = 0.05/12$, with 12 comparisons per behavioural rating).

	Amount of movement			Fast movement			Vertical movement		
	r	p-value	Corr. p-value	r	p-value	Corr. p-value	r	p-value	Corr. p-value
Fast movement	.797	.000	.000						
Vertical movement	.472	.000	.000	.487	.000	.000			
Forward-away	.044	.086	.999	.003	.916	.999	-.014	.589	.999

Abbreviation: Corr.: Bonferroni-corrected p-value

Table SR3. Correlation between representational dissimilarity matrices of the behavioural ratings and the computed features. The correlations were computed using Spearman's rank correlation. Both uncorrected and Bonferroni-corrected p-values are displayed for each comparison ($\alpha_{\text{bonf}} = 0.05/12$, with 12 comparisons per behavioural rating).

		Computed features														
		Emotional category			Velocity			Acceleration			Vertical movement			Limb angles		
		r	p-value	Corr.	r	p-value	Corr.	r	p-value	Corr.	r	p-value	Corr.	r	p-value	Corr.
				p-value			p-value			p-value			p-value			p-value
Behavioural ratings	Emotional category	1.000	.000	.000	.094	.000	.003	.101	.000	.001	.057	.025	.302	.251	.000	.000
	Emotional rating	.728	.000	.000	.174	.000	.000	.147	.000	.000	.053	.037	.445	.280	.000	.000
	Emotional intensity	.414	.000	.000	.100	.000	.001	.150	.000	.000	-.079	.002	.023	.251	.000	.000
	Valence	.466	.000	.000	.226	.000	.000	.112	.000	.000	.0134	.000	.000	.121	.000	.000
	Action rating	.720	.000	.000	.123	.000	.000	.076	.003	.032	.065	.010	.123	.339	.000	.000
	Familiarity	.337	.000	.000	.018	.479	.999	.086	.001	.009	-.107	.000	.000	.133	.000	.000
	Contraction	.414	.000	.000	.075	.003	.038	.152	.000	.000	-.061	.016	.191	.325	.000	.000
	Symmetry	.268	.000	.000	.109	.000	.000	.063	.013	.158	-.019	.446	.999	.242	.000	.000
	Amount of movement	.323	.000	.000	.340	.000	.000	.319	.000	.000	.015	.561	.999	.202	.000	.000
	Fast movement	.270	.000	.000	.441	.000	.000	.423	.000	.000	.026	.311	.999	.109	.000	.000
	Vertical movement	.213	.000	.000	.555	.000	.000	.334	.000	.000	.328	.000	.000	.010	.692	.999
	Forward-away	.535	.000	.000	.035	.170	.999	.077	.003	.031	.147	.000	.000	.160	.000	.000

Abbreviation: Corr.: Bonferroni-corrected p-value

Table SR3 (continuation). Correlation between representational dissimilarity matrices of the behavioural ratings and the computed features. The correlations were computed using Spearman's rank correlation. Both uncorrected and Bonferroni-corrected p-values are displayed for each comparison ($\alpha_{\text{bonf}} = 0.05/12$, with 12 comparisons per behavioural rating).

		Computed features											
		Symmetry			Shoulder ratio			Surface			Limb contraction		
		r	p-value	Corr.	r	p-value	Corr.	r	p-value	Corr.	r	p-value	Corr.
				p-value			p-value			p-value			p-value
Behavioural ratings	Emotional category	.262	.000	.000	.185	.000	.000	.036	.161	.999	.112	.000	.000
	Emotional rating	.300	.000	.000	.220	.000	.000	.075	.003	.039	.119	.000	.000
	Emotional intensity	.400	.000	.000	.356	.000	.000	-.024	.343	.999	.113	.000	.000
	Valence	.049	.057	.684	.279	.000	.000	.218	.000	.000	.003	.915	.999
	Action rating	.365	.000	.000	.296	.000	.000	.136	.000	.000	.123	.000	.000
	Familiarity	.155	.000	.000	.207	.000	.000	-.018	.470	.999	.103	.000	.001
	Contraction	.332	.000	.000	.147	.000	.000	.005	.853	.999	.329	.000	.000
	Symmetry	.366	.000	.000	.330	.000	.000	.178	.000	.000	.049	.054	.648
	Amount of movement	.343	.000	.000	.233	.000	.000	-.014	.589	.999	.109	.000	.000
	Fast movement	.224	.000	.000	.157	.000	.000	-.048	.061	.736	.067	.009	.108
	Vertical movement	.013	.602	.999	.136	.000	.000	.028	.264	.999	-.021	.420	.999
	Forward-away	.096	.000	.002	.000	.997	.999	-.027	.284	.999	.180	.000	.000

Abbreviation: Corr.: Bonferroni-corrected p-value

References

- Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7291-7299.
- Kleinsmith, A., & Bianchi-Berthouze, N. (2012). Affective body expression perception and recognition: A survey. *IEEE Trans. Affect. Comput.*, 4(1), 15-33.