The Body Action and Posture Coding System (BAP): Development and Reliability

Nele Dael · Marcello Mortillaro · Klaus R. Scherer

Abstract Several methods are available for coding body movement in nonverbal behavior research, but there is no consensus on a reliable coding system that can be used for the study of emotion expression. Adopting an integrative approach, we developed a new method, the body action and posture coding system, for the time-aligned micro description of body movement on an anatomical level (different articulations of body parts), a form level (direction and orientation of movement), and a functional level (communicative and self-regulatory functions). We applied the system to a new corpus of acted emotion portrayals, examined its comprehensiveness and demonstrated intercoder reliability at three levels: (a) occurrence, (b) temporal precision, and (c) segmentation. We discuss issues for further validation and propose some research applications.

Keywords Body movement · Expression · Coding · Method · Emotion

Introduction

Advancement of knowledge in an empirical research area often goes hand in hand with methodological development. In the affective sciences, more specifically in the field of emotion communication, research on facial and vocal emotion expression and perception has rapidly progressed after the establishment of facial movement scoring techniques

---

N. Dael · M. Mortillaro · K. R. Scherer
Swiss Centre for Affective Sciences, Faculty of Psychology and Educational Sciences,
University of Geneva, Geneva, Switzerland

N. Dael (✉)
Institute of Psychology, University of Lausanne, Bâtiment Anthropole, Quartier Dorigny,
1015 Lausanne, Switzerland
e-mail: Nele.Dael@unil.ch

(Ekman and Friesen 1976, 1978; Ekman et al. 1971; Izard 1979, 1983), and refined techniques for speech analysis (Juslin and Scherer 2005; Scherer 1979, 1982; Scherer et al. 1991). A similar level of methodological advancement in the field of body movement research has been largely absent. A number of methods exist for coding body movement, but there is currently no consensus on what the basic unit of body movement is for analysis or on gesture categorization, nor is there a theoretical framework for understanding body movement (Harrigan 2005). The overall aim of this paper is to describe the development of a coding system, capable of providing systematic and reliable descriptions of body movements with a particular focus on bodily expression of emotion.

Techniques for Measuring Body Movement in Emotion Expression Research

Several kinds of measurement exist for investigating the bodily expression of emotion. These measurements can be organized along a continuum based on the degree of subjective inference required from the observer (Scherer and Wallbott 1985). At one end of the continuum lies subjective judgment of movement (e.g., Arieti 1959; James 1932; Wolff 1951). Skilled researchers and clinicians described behavior on the basis of their inferences and intuitions without using systematic labels. These studies produced interesting findings but they did not capture objective movement characteristics, which can be replicated by other researchers and thus be used as valid indicators of the expressed emotion.

Measures that do not require any human observer can be situated on the other end of the continuum. Here one finds techniques that directly measure the production process at the muscular level. Motor variables are obtained by attaching physiological instruments to the moving body (e.g., electromyograph, accelerometer). These systems provide very accurate indicators of motoric movement (e.g., Coombes et al. 2007), but they can be used only in strictly controlled laboratory settings, they are obtrusive, and often limit behavioral freedom.

Motion tracking systems are another example of techniques providing objective measures of body movement, for example gait (Omlor and Giese 2006), and emotional arm movements (Pollick et al. 2001). Motion tracking also requires a very controlled environment and specialized cameras or sensors to detect positions of light-reflecting patches attached to the body joints.

Video based automatic tracking techniques are a less obtrusive way of measuring body movement. Such systems rely on color or silhouette extraction algorithms and are successfully used in the area of affective computing (e.g., Glowinski et al. 2011). These systems can detect gross body movements but they do not provide fine-grained description of the movements of specific body parts.

Observational coding occupies the middle position in terms of time- and cost-efficiency compared to the above-mentioned techniques of movement description. Observational coding is based on explicit operational definitions of an a priori defined set of behavior codes and follows a fixed coding procedure. Codes and procedures are part of a coding system, i.e., a standard protocol for describing behavior (Bakeman and Gottman 1987; Scherer and Wallbott 1985). Unlike subjective behavior judgments, descriptions based on observational coding can provide a detailed account of body movement with less degree for observer inference. Ideally, the resulting variables refer to movement characteristics that are actually present and thus can be verified and replicated by another coder. Hence, much depends on the reliability and applicability of the coding system.

Coding Systems for Studying Nonverbal Emotion Communication: Movement Quality or Movement Type

A diverse range of coding systems has been produced over the years (for a detailed overview of methodologies for coding body movement, see Harrigan 2005; Rosenfeld 1982). In general, two approaches to observational coding can be distinguished; one approach focuses on movement quality, the other on movement type.

In the first approach, researchers focus on the way in which movements are executed with respect to the dimensions of space and time. They define movement quality descriptors such as speed, smoothness, tension, and force. In the study of dance, this approach is exemplified by Labanotation (Hutchinson 1961; Laban 1956, 1975). Modeled after musical notation systems, symbols are used to represent the actions or positions of the body (for a review and illustration see Harrigan 2005; Rosenfeld 1982). The qualitative aspects of body movement are represented by the general components of effort and shape (e.g., fluency, strength, weakness), which seem closely related to emotion as shown in several perception studies (de Meijer 1989; Gross et al. 2010). However, expert training is very time-consuming. Furthermore, the coding reliability remains undocumented and many of the definitions of the qualitative terms have strong subjective connotations making them prone to observer bias (Rosenfeld 1982).

The Bernese system is another example of the movement quality approach based on spatio-temporal coding (Frey and Pool 1976; Frey and von Cranach 1973; see Harrigan 2005; Rosenfeld 1982). This coding system provides a detailed and highly reliable description of the positions and movement of body parts along the three Cartesian axes (sagittal, vertical, and transverse). On the basis of elementary spatio-temporal properties, Frey and colleagues formulated objective indexes of movement qualities such as openness, expansion, and reaching out. The Bernese coding system has been applied and further extended in the field of computer animations (e.g., Bente et al. 2001), but has rarely been used in the study of nonverbal behavior and emotion psychology. On one hand, this coding system proved to be accurate and objective. On the other hand, training and coding is very time-consuming and needs considerable post-processing, thus limiting its applicability to studies with a large number of participants.

In the second approach to observational coding, scholars propagate a typology or classification of body movement. In these cases, coding systems are often tailored to investigate specific research purposes, and consequently they differ on the level of description (anatomical articulation, movement form, or function) and on the body parts or subsystems that are distinguished (actions, gestures, or postures). For example, Birdwhistell (1970) created a fine-grained coding system based on linguistic principles, but its use in psychological studies has been limited by the lack of supporting validation and by the controversial structural analogy between body motion and speech (for a review see Harrigan 2005; Rosenfeld 1982).

Different coding systems have been designed for studying specific topics in psychology. Mehrabian (1972) proposed a coding system limited to body orientation and posture for studying interpersonal attitudes. Bull (1987) devised an elaborate set of systems to describe in detail the form and orientation of postures and actions during seated interaction to investigate communication of attitudes and emotions (boredom/interest, agreement/disagreement). Recently, in the framework of their work on the expression of pride and other self-conscious emotions, Tracy and Robins (2007) have developed a short nonverbal coding scheme for assessing prototypical components of pride in static upper body posture. One of the most influential contributions is the typology of hand actions developed in the

context of bodily communication of emotion and deception by Ekman and Friesen (1969, 1972). Their typology has later been expanded into a functional linguistic classification system of hand actions accompanying speech (defined as gestures) that is included in a widely used gesture-speech coding system (Duncan et al. 1995; McNeill 2005). In this system, the function of a gesture is inferred from the verbal content in gesture accompanying speech, so the resulting descriptions do not refer to movement properties alone.

Others have attempted to integrate different levels and approaches. Birdwhistell's structural coding system in fact also included some movement quality descriptors, called "motion qualifiers" such as intensity and range (Birdwhistell 1970). Wallbott (1998) proposed a selective but reliable collection of both form and functional categories from different body parts and subsystems, and even included movement qualities such as movement activity, expansiveness, and energy. The category selection was based on free format descriptions of the material and has not yet been replicated in other studies. However, this integrative (yet not comprehensive) approach has proven effective in the analysis of emotion expression (Wallbott 1998).

Segmentation of the Behavior Stream

Behavior segmentation or unit definition is a major prerequisite to analyze individual behaviors and to investigate patterns of behaviors. Time-stamped behavior units further allow the study of timing of units or even subunits in relationship to other co-occurring internal and external events such as facial configurations or the response of an interaction partner. The Facial Action Coding System (FACS) includes explicit rules for coding unit boundaries and phases (onset, apex, and offset) of facial muscle movement (Ekman and Friesen 1978). Such rules are rarely specified in body movement coding systems (for a notable exception see Birdwhistell 1970). Conversely, segmentation has been investigated intensively in cognitive linguistics as researchers in this field are interested in the synchronization between speech and gesture (McNeill 2005).

Within cognitive linguistics a consensus has been reached on a set of syntactic rules for segmenting gestures in at least three phases: preparation, stroke, and retraction. During the preparation phase the limb moves away from the resting position into the gesture space where it can begin the stroke (for example lifting the arm in front of the chest). The stroke is defined as the most energetic phase of the gesture that carries meaning linked to the lexical affiliate (a word or phrase in co-occurring speech—for example, showing an "ok" sign). Retraction is the phase in which the body part returns to a resting position (for example dropping the arm to the side the body). Gesture segmentation strongly relies on semantic interpretation of the movement derived from the frame of discourse. Consequently, one cannot identify the stroke phase and therefore any linked segment without information from the semantic framework (see appended Duncan Coding Manual in McNeill 2005). These gesture segmentation rules are furthermore not generalizable to body actions and postures that do not have any verbalized meaning. Taken together, no validated system for defining and segmenting movement units is currently available for body movement research in general and emotion expression research in particular.

The Present Research

The preceding literature review shows that coding systems for describing body movement are very diverse and spread across different disciplines. Few have been successfully applied in the study of nonverbal emotion expression. Therefore, our goal was to construct

a comprehensive coding system for body movement that can be used as a research tool in emotion expression research, and that is extendable to other areas of nonverbal research. In this study we describe the development of the body action and posture (BAP) coding system and present the first validation study on a representative set of acted nonverbal emotional portrayals.

Method

The GEMEP Corpus as Development Material

This coding system was developed using the Geneva Multimodal Emotion Portrayals (GEMEP) corpus (Bänziger and Scherer 2010). For this study we adopted a varied, representative set of emotional expressions that has been rated as recognizable and believable (Bänziger et al. 2011; Bänziger and Scherer 2010). This material consists of 151 portrayals of 17 emotions encoded by 10 actors (5 male). The emotional states range from modal emotions to more subtle emotions, and are balanced in terms of the two major emotion dimensions; valence and arousal. The actors received short descriptions of the emotional states and encoded the intended emotion within a scenario-based social interaction setting, while uttering one of two standard pseudo-sentences ("ne kali bam soud molen", "koun se mina lod belam"). The actors did not receive instructions to use particular facial or bodily expressions, apart from the restriction not to move away from the two cameras. Recordings were made in a controlled environment under high technical standards in terms of frame speed and multiple camera angles (information from three dimensions). Two camera angles captured the body from the knees upwards: The frontal view represents the perspective of an interlocutor, the side view represents the perspective of an observer standing to the right of the actor. The total duration of this dataset is 9,433 frames recorded at 25 FPS (frames per second). The average duration of a video portrayal is 62.47 frames (SD = 25.44), which corresponds to two and a half seconds. We refer to Bänziger and Scherer (2010) and Bänziger et al. (2011) for a full description of the material, procedure and rationale.

Description of the BAP Coding System

Here we describe the general framework on which the BAP coding system is built. The main distinction is made between body posture units and body action units (Harrigan 2005). A posture unit represents the general alignment of one or a set of articulators (head, trunk, arms) to a particular resting configuration, which shows periodic changes known as posture shifts (e.g., a person leaning backward, arms crossed). An action unit is a local excursion of one or a set of articulators (mostly the arms) outside a resting configuration with a very discrete onset (start point), a relatively short duration, and a distinct offset (end point) where the articulator returns to a resting configuration (e.g., head shake, pointing arm gesture). All visible skeletal body movement was categorized as an action or a posture, with the exception of leg movement for which there was not enough visibility to distinguish between action and posture.

Codes describing postures and actions are organized on two integrated levels. On a first level, BAP codes describe *anatomical articulation* (which part of the body is actively moving) following well-established kinesiological standards (Hinson 1977; Neumann 2002). The BAP coding system allows the coding of articulation of the neck (head), trunk

(including spine and chest), upper and lower arms (shoulders, elbows, wrists, and fingers), and the lower limbs (knees). We further included general categories such as leg movement and whole body movement for which the exact articulators could not be obtained due to technical limitations. Importantly, by focusing on articulation, BAP codes apply only to spatial movement caused by muscle contraction and not to passive displacement of any connected articulator. For example, spatial displacement of the head caused by trunk bending is passive and thus only articulation of the trunk is coded.

On a second level, BAP codes refer to the *form of the movement* of the active body part (how the part of the body is moving). Movement is described in terms of direction and orientation in three-dimensional space by adopting an anatomical and an external reference frame (Rosenfeld 1982). Movement direction was defined with respect to three orthogonal axes of the body (the sagittal, vertical, and transverse axis). The direction of a posture unit (e.g., head to left) refers to the end position compared to the anatomical standard position. On the other hand, the direction of an action unit is coded independently and refers to the movement itself. Orientation is coded continuously from the viewpoint of an interlocutor standing in front of the actor. The active body part was coded as facing or averted from the interlocutor's point of view.

Actions are additionally described on a *functional level* that is parallel to the anatomical and form description level. Functional BAP codes include the three behavioral classes identified by Ekman and Friesen (1972): emblems, illustrators, and manipulators. Emblems are conventionalized actions that have precise meanings. These symbolic actions are culturally defined and can be used independently of speech. Because all GEMEP actors belong to the French speaking community, we adopted a descriptive semiotic lexicon on French gestures as a guide to code emblematic action (Calbris and Doyle 1990, pp. 2–7, Table 1 and illustrations). Manipulators are actions in which one part of the body manipulates another body part or an object for some type of body contact, such as scratching, rubbing, etc. We distinguished self- and object manipulators and coded both the articulating body part (called adaptor) and the body part or clothes being manipulated or adapted. Illustrators are conversational actions that support accompanying speech. They illustrate the rhythm or content of a verbalized message. Contrary to emblems they are tied to speech and do not have a fixed meaning. We selected two illustrator subtypes that can possibly be detected independently of semantic content from the speech context; beat and deictic. Beats are repetitive actions that accentuate points in time, illustrating structural or rhythmic aspects of co-occurring speech. A deictic is a referential action indicating a real or abstract object, person, event or location in space (e.g., pointing).

It has been argued that these types of actions are often multifaceted in that several functional or semiotic properties can be simultaneously present (McNeill 2005, pp. 38–41). Functional categories are thus not mutually exclusive but can be conceptualized as dimensions on which a behavior has certain loadings (e.g., beat-like deictic pointing). We adapted this dimensional approach to the functional categories in our coding system by applying ordinal coding on the saliency of the behavior in terms of duration, location, and execution. A functional unit is coded as *very subtle* (1) when it is executed with minor articulation (e.g., deictic pointing with a partly extended finger), for a very short period of time (e.g., few repetitions of a beat), or located in the peripheral movement space of the person (e.g., below the hips). On the other extreme, a functional unit is coded as *very pronounced* (5) when it is supported by maximum articulation (e.g., deicitic pointing with fully extended finger and arm), stretched in time over several seconds (many repetitions of a beat), or located in the central movement space of the person (e.g., near the face).

Table 1 Complete list of behaviors coded with the BAP coding system

	Behavior variable	Short description
Head orientation		
1	Facing	The face is oriented towards the interlocutor
2	Averted	The face is oriented away from the interlocutor
Head posture		
3	Lateral head turn towards a left position	A rotation of the head around the vertical axis that results in the face becoming more visible on right side relative to the anatomical standard position
4	Lateral head turn towards a right position	A rotation of the head around the vertical axis that results in the face becoming more visible on left side relative to the anatomical standard position
5	Lateral head turn towards the lateral middle position	A rotation of the head around the vertical axis that results in the anatomical standard position
6	Lateral head tilt towards a left position	A rotation of the head around the sagittal axis that results in the head tilted laterally towards the left shoulder
7	Lateral head tilt towards a right position	A rotation of the head around the sagittal axis that results in the head tilted laterally towards the right shoulder
8	Lateral head tilt towards the lateral middle position	A rotation of the head around the sagittal axis that results in the anatomical standard position
9	Vertical head tilt towards an upward position	A rotation of the head around the transversal axis that results in the head lifted up relative to the anatomical standard position
10	Vertical head tilt towards a downward position	A rotation of the head around the transversal axis that results in the head dropped down relative to the anatomical standard position
11	Vertical head tilt towards the vertical middle position	A rotation of the head around the transversal axis that results in the anatomical standard position
12	Neck extension towards a forward head position	A translatory motion of the head along the sagittal axis that results in the head put forward relative to the anatomical standard position
13	Neck retraction towards a backward head position	A translatory motion of the head along the sagittal axis that results in the head put backward relative to the anatomical standard position
14	Neck extension/retraction towards the frontal middle position	A translatory motion of the head along the sagittal axis that results in the anatomical standard position
Trunk orientation		
15	Facing	The trunk is oriented towards the interlocutor
16	Averted	The trunk is oriented away from the interlocutor
Trunk posture		
17	Trunk lean towards a forward position	The upper body leans forward relative to the anatomical standard position
18	Trunk lean towards a backward position	The upper body leans backward relative to the anatomical standard position
19	Trunk lean towards the frontal middle position	The upper body moves along the sagittal axis (forward or backward) that results in the anatomical standard position
20	Lateral trunk lean towards a left position	A lateral flexion of the trunk that results in the upper body leaning laterally to the left relative to the anatomical standard position

Table 1 continued

	Behavior variable	Short description
21	Lateral trunk lean towards a right position	A lateral flexion of the trunk that results in the upper body leaning laterally to the right relative to the anatomical standard position
22	Lateral trunk lean towards the lateral middle position	A lateral flexion of the trunk (left or right) that results in the anatomical standard position
23	Lateral trunk rotation towards a left position	A rotation of the trunk around the vertical axis that results in a left alignment of the upper body relative to the anatomical standard position
24	Lateral trunk rotation towards a right position	A rotation of the trunk around the vertical axis that results in a right alignment of the upper body relative to the anatomical standard position
25	Lateral trunk rotation towards the lateral middle position	A rotation of the trunk around the vertical axis that results in the standard anatomical position
26	Spine movement towards an erected position[b]	The spine is straightened so the trunk is more erect or upright relative to the anatomical standard position
27	Spine movement towards a bent position[b]	The spine is bent so the trunk is less erect or upright relative to the anatomical standard position
28	Spine movement towards the standard position[b]	The spine takes a normal curvature associated to the anatomical standard position
29	Chest movement towards an upward or forward position[b]	The chest is lifted up or turned outward relative to the anatomical standard position
30	Chest movement towards a downward or inward position[b]	The chest is lowered or turned inward relative to the anatomical standard position
Whole body posture		
31	Whole body moves or leans towards a forward position	The whole body, including the lower part (hips and legs) moves or leans forward relative to the anatomical standard position
32	Whole body moves or leans towards a backward position	The whole body, including the lower part (hips and legs) moves or leans backward relative to the anatomical standard position
33	Whole body moves or leans towards the frontal middle position	The whole body, including the lower part (hips and legs) moves or leans forward or backward towards the standard anatomical position
34	Whole body moves or leans towards a left position	The whole body, including the lower part (hips and legs) moves or leans towards the left relative to the anatomical standard position
35	Whole body moves or leans towards a right position	The whole body, including the lower part (hips and legs) moves or leans towards the right relative to the anatomical standard position
36	Whole body moves or leans towards the lateral middle position	The whole body, including the lower part (hips and legs) moves or leans left or right towards the standard anatomical position
Arms posture		
37	Left arm at side	The left arm and hand hangs at the side of the body
38	Left arm held in front	The left hand and/or arm is held in front of the body
39	Left arm held behind back[b]	The left hand and/or arm is held behind the back
40	Left hand in pocket	The left hand (or a part of it) is in the pocket
41	Left hand at waist	The left hand rests on the hip
42	Left hand at back of head/neck[a]	The left hand is held behind the back of the head or neck

Table 1 continued

	Behavior variable	Short description
43	Right arm at side	The right arm and hand hangs at the side of the body
44	Right arm held in front	The right hand and/or arm is held in front of the body
45	Right arm held behind back[b]	The right hand and/or arm is held behind the back
46	Right hand in pocket	The right hand (or a part of it) is in the pocket
47	Right hand at waist	The right hand rests on the hip
48	Right hand at back of head/neck[a]	The right hand is held behind the back of the head or neck
49	Arms crossed	The arms are crossed in front of the body
50	Both arms held in front	Both hands and/or arms hold each other in front of the body
51	One arm holds other in front	One arm rests on the other or is held by the other in front of the body
52	One arm holds other in back	One arm rests on the other or is held by the other behind the back
53	Left shoulder up[a]	The left shoulder is lifted up relative to the anatomical standard position
54	Left shoulder down[a]	The left shoulder is dropped down, lowered to the ground relative to the anatomical standard position
55	Left shoulder to the front[a]	The left shoulder is put to the front relative to the anatomical standard position
56	Left shoulder to the back[a]	The left shoulder is pulled back relative to the anatomical standard position
57	Right shoulder up[a]	The right shoulder is lifted up relative to the anatomical standard position
58	Right shoulder down[a]	The right shoulder is dropped down, lowered to the ground relative to the anatomical standard position
59	Right shoulder to the front[a]	The right shoulder is put to the front relative to the anatomical standard position
60	Right shoulder to the back[a]	The right shoulder is pulled back relative to the anatomical standard position
61	Symmetrical arms posture	The postural alignment of the arms is symmetrical
62	Asymmetrical arms posture	The postural alignment of the arms is asymmetrical
Gaze		
63	Toward	The gaze is directed towards the interlocutor
64	Upward	The gaze is directed forward and above the interlocutor
65	Downward	The gaze is directed forward and below the interlocutor
66	Averted sideways	The gaze is directed away from the interlocutor to the left or right, regardless of vertical direction (up or down)
67	Eyes closed	Both eyes are closed
Head action		
68	Left head turn	A left rotation of the head around the vertical axis
69	Right head turn	A right rotation of the head around the vertical axis
70	Left head tilt	A left rotation of the head around the sagittal axis
71	Right head tilt	A right rotation of the head around the sagittal axis
72	Upward head tilt[a]	An upward rotation of the head around the transversal axis
73	Downward head tilt[a]	A downward rotation of the head around the transversal axis

Table 1 continued

	Behavior variable	Short description
74	Forward neck extension[a]	A forward translatory motion of the head along the sagittal axis
75	Backward neck retraction[a]	A backward translatory motion of the head along the sagittal axis
76	Up-down head shake	The head is repeatedly tilted up and down
77	Left–right head shake	The head is repeatedly rotated or tilted left and right
78	Forward–backward head shake	The neck is repeatedly extended and retracted forward and backward
79	Head action hold	The head is held static while performing an action
Trunk action		
81	Forward trunk lean	A forward leaning movement of the trunk
82	Backward trunk lean	A backward leaning movement of the trunk
83	Left trunk lean[a]	A lateral flexion of the trunk to the left
84	Right trunk lean[a]	A lateral flexion of the trunk to the right
85	Left trunk turn[a]	A left rotation of the trunk around the vertical axis
86	Right trunk turn[a]	A right rotation of the trunk around the vertical axis
87	Spine straightening[a]	The spine is erected
88	Spine bending[a]	The spine is relaxed
89	Upward/forward chest movement[a]	The chest is lifted up or turned outward
90	Downward/inward chest movement[a]	The chest is lowered or turned inward
91	Up-down chest movement[a]	The chest repeatedly moves up/forward and down/inward
92	Left–right trunk leaning[a]	The trunk repeatedly leans to the left and the right
93	Left–right trunk rotation[a]	The trunk repeatedly rotates to the left and the right
94	Forward–backward trunk leaning[a]	The trunk repeatedly leans forward and backward
95	Trunk action hold[a]	The trunk is held static while performing an action
Arm action		
96	Left finger(s) articulation	Action articulation of one or more fingers of the left hand
97	Left wrist articulation	Action articulation of the left wrist resulting in movement of the whole hand
98	Left elbow articulation	Action articulation of the left elbow resulting in movement of the forearm
99	Left shoulder articulation	Action articulation of the left shoulder resulting in movement of the upper arm
100	Left arm action upward	The left arm moves upward
101	Left arm action downward	The left arm moves downward
102	Left arm action forward	The left arm moves forward
103	Left arm action backward	The left arm moves backward
104	Left arm action to the left	The left arm moves to the left
105	Left arm action to the right	The left arm moves to the right
106	Left arm action away from the body	The left arm moves away from the body
107	Left arm action towards the body	The left arm moves towards the body
108	Left arm action vertical repetition	The left arm repeatedly moves up and down

Table 1 continued

	Behavior variable	Short description
109	Left arm action lateral repetition	The left arm repeatedly moves to the left and right
110	Left arm action frontal repetition	The left arm repeatedly moves forward and backward
111	Left arm action curved repetition[a]	The left arm repeatedly moves in a circular fashion
112	Left arm action hold	The left arm is held static while performing an action
113	Left arm action retraction	The left arm moves back to a resting position
114	Right finger(s) articulation	Action articulation of one or more fingers of the right hand
115	Right wrist articulation	Action articulation of the right wrist resulting in movement of the whole hand
116	Right elbow articulation	Action articulation of the right elbow resulting in movement of the forearm
117	Right shoulder articulation	Action articulation of the right shoulder resulting in movement of the upper arm
118	Right arm action upward	The right arm moves upward
119	Right arm action downward	The right arm moves downward
120	Right arm action forward	The right arm moves forward
121	Right arm action backward	The right arm moves backward
122	Right arm action to the left	The right arm moves to the left
123	Right arm action to the right	The right arm moves to the right
124	Right arm action away from the body	The right arm moves away from the body
125	Right arm action towards the body	The right arm moves towards the body
126	Right arm action vertical repetition	The right arm repeatedly moves up and down
127	Right arm action lateral repetition[b]	The right arm repeatedly moves to the left and right
128	Right arm action frontal repetition	The right arm repeatedly moves forward and backward
129	Right arm action curved repetition[a]	The right arm repeatedly moves in a circular fashion
130	Right arm action hold	The right arm is held static while performing an action
131	Right arm action retraction	The right arm moves back to a resting position
132	Symmetrical arms action	Both arms jointly move in a symmetrical fashion
133	Asymmetrical arms action	Both arms jointly move in an asymmetrical fashion
Other		
134	Touch	One body part touches another body part or an object as part of an action
135	Knee bend	One knee or both knees are bent or bending
136	Leg movement	Any movement of the lower limbs
Action function		
137	Emblem	A symbolic and conventionalized body action with a culturally defined fixed form-meaning relationship
138	Illustrator	A conversational action that supports accompanying speech by illustrating the rhythm or content of a verbalized message
139	Beat	Repetitive action that accentuates points in time, illustrating structural or rhythmic aspects of co-occurring speech (type of illustrator)
140	Deictic	A referential action indicating a real or abstract object, person, event or location in space (type of illustrator)

Table 1 continued

	Behavior variable	Short description
141	Manipulator	An action in which one part of the body manipulates another body part or an object for some type of body contact

Full descriptions of the behavior variables are provided in the BAP manual included in the Electronic supplementary Material

[a] Variable removed due to non-occurrence (≤.01)

[b] Variable removed because the occurrence rate of the pooled variable to which the variable belongs (see Appendix C in the Electronic Supplementary Material) was too low to determine reliability (≤.05)

BAP units are time-locked and represent temporal behavioral segments (i.e., with a beginning and an end). The definitions of all categories thus include criteria for onset and offset. On the basis of these criteria, a behavioral event is described through one or multiple units that are not mutually exclusive: they can overlap in time (e.g., left head turn + head tilt down) and belong to multiple descriptive levels simultaneously (e.g., forward arm action + deictic).

Furthermore, posture units are segmented into a configuration phase (time segment where the end position of the coded articulator is maintained for the coded direction) and a transition phase ("postural shift", i.e., the time segment where the articulator performs the movements for obtaining the end position). Action units are segmented into subunits according to differences in the form properties of the movement. An action is segmented into a sequence of discrete subunits that are separated by discontinuous changes of direction (based on Kita et al. 1998). Repetitive movement is not segmented or coded for each repeated cycle separately. Units identified by functional codes are not further segmented.

In addition to these three levels describing skeletal body movement, the BAP coding system also includes categories for coding gaze direction from the viewpoint of an interlocutor standing in front of the actor.

Table 1 presents a list of all 141 behaviors (actions, postures, and other behavior categories) included in the BAP coding system including short definitions. Full descriptions can be found in the coding manual that is available as Electronic Supplementary Material.[1] The manual is accompanied by a written document stating additional coding guidelines.

Coding Procedure

We used the Anvil software (Kipp 2001, 2003, 2007; Kipp et al. 2007) for video rendering and time-locked coding. We specified the coding scheme and manual in XML format. Anvil implements the coding scheme into an intuitive graphical user interface through which the coder can assign codes to the body movements specifying the beginning and the end of the behaviors (Fig. 1). Anvil allows the coder to view the video frame by frame or at different speeds and at the same time offers the possibility to browse the HTML coding manual any time during the coding.

---

[1] Interested researchers can obtain the coding manual with coding guidelines at the online journal's website or at http://www.affective-sciences.org/gemep/BAP_coding
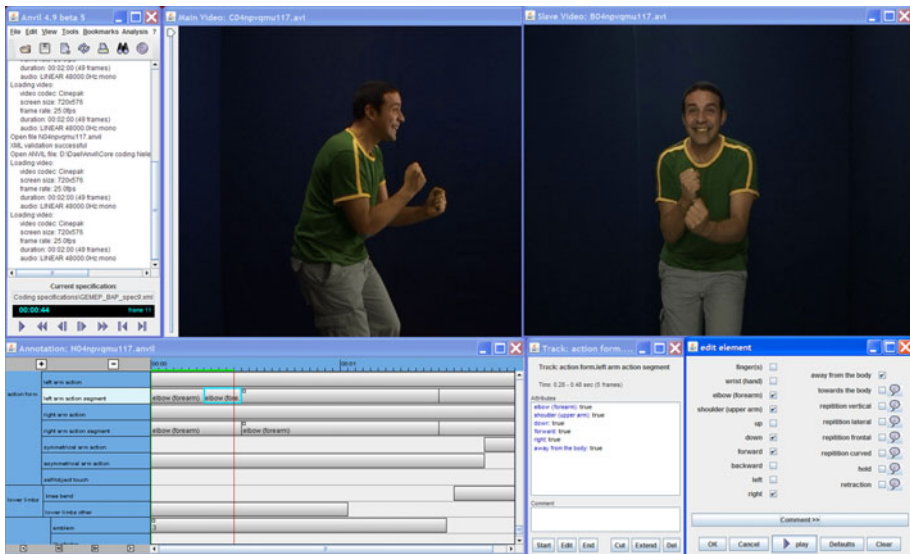
Fig. 1 The coding platform: customized graphical user interface in Anvil

Two trained observers[2] independently coded the GEMEP dataset using the Anvil software according to the coding definitions and guidelines described in the manual and supporting documentation. The final coding period -approximately 38 h for the entire dataset, i.e., 15 min per portrayal—was preceded by several training sessions[3] including discussion and further clarification of the coding manual. This made both coders well acquainted with the kind of material, coding procedure, and the software. These trials also served as feasibility checks for the coding system's implementation within the Anvil software module.

We encrypted the emotion label from the filenames to ensure that the coders were blind to the emotion portrayed by the actors. Also, the speech sound was turned off so that the coding was informed by visual movement only. Both videos of the frontal and profile view were simultaneously present during the coding, and could be played in synchronized fashion and at reduced speed.

The coders first viewed each portrayal at normal or reduced speed to detect behavior occurrences in terms of visual movement and static postures, then manually scrolled through the video frame by frame to code the onsets and offsets of the present units and subunits (segments). They followed a fixed order of coding (see guidelines in the Electronic Supplementary Material), starting with focusing on a particular body part, then distinguishing active from passive movement and body posture from body action, followed by coding the behavior units along the descriptive levels defined above (articulation, movement form, and movement function).

---

[2] One coder (main investigator, hence called the first coder) had some prior experience in behavior observation whereas the other coder (research assistant, hence called the second coder) had no experience in coding behavior prior to this study.

[3] The training material consisted of 31 portrayals that were taken from the GEMEP corpus. These are *not* included in the material that is presented in this paper to test reliability.

Data Analysis

The output data files used for analysis are text files where each frame of a portrayal is represented in a row and the behavior variables (see Table 1) are represented in columns. For all variables except the functional codes, the values are 0 for *absence* and 1 for *presence*. For the functional categories, the values are between 0 and 5. We assessed intercoder reliability at three levels of analysis: (1) occurrence or non-occurrence within the portrayal; (2) temporal precision; and (3) segmentation.

We assessed occurrence agreement by computing Cohen's kappa (Cohen 1960; Fleiss 1981) based on the presence or absence of a behavior unit within the portrayal ($N = 151$). This statistic is a widely used yet not unequivocal intercoder agreement measure for binary and nominal coding. Kappa is preferable over simple percentage of agreement because it corrects for agreement due to chance. We chose kappa also because we wanted to credit both agreement on presence and absence of a category but avoid agreement inflation due to high numbers of non-occurrences for some behaviors.

Bakeman and Gottman (1987) already highlighted the problem of assessing agreement on a temporal basis where the behavioral stream is differently segmented into successive events by two or more coders. Therefore, to assess reliability of temporal precision we followed their recommendation to compute kappa on time intervals. This approach was also adopted in a psychometric evaluation of FACS coding (Sayette et al. 2001). As the coding was performed at a very high precision level (coders watched the videos at highly reduced speed, often manually moving up and down frames) we first computed the kappa at the smallest interval, namely presence or absence within the same frame for both coders (40 ms or 1/25 s, $N = 9,433$). Next, we applied two tolerance levels and recomputed kappa using a dynamic time-window, where the value (presence or absence) at each frame from the second coder (index frame) was compared to the values of an indexed range of frames from the first coder. The range of the first tolerance level was five frames, i.e., two frames before and after the index frame (200 ms or 5/25 s). The range of the second tolerance level was eleven frames, i.e., five frames before and after the index frame (440 ms or 11/25 s). These time-based kappa computations preserve a maximum of information coded, notably on timing and duration, and allow comparison of different levels of precision.

We measured agreement on the segmentation of postures in transition and configuration phases by computing the kappa on portrayals where both coders indicated the presence of the posture. To assess action segmentation agreement we calculated the number of subunits coded by each coder in every portrayal, and then correlated these frequencies over all portrayals. High correlation indicates that there is a linear relationship between the frequencies of subunits coded by both coders. Temporal precision of action segmentation could not be computed on the basis of our data-output for technical reasons.

Interpretation of the coefficients is guided by Fleiss (1981), who considers a kappa between .40 and .60 as fair, between .60 and .75 as good and over .75 as excellent. We also follow Bakeman and Gottman (1987) who take a kappa of less than .70 with some concern. We refrained from interpreting kappas of behaviors that were coded in less than 5% of the portrayals or of the total duration of the dataset because they do not accurately reflect reliability.

In an effort to summarize information and increase stability of estimate, we followed a sequential pooling approach starting from the elementary variable corresponding to the basic movement unit (e.g., head turn to left, head turn to right) and selectively grouped units that belong to the same underlying behavioral category (e.g., lateral head turn) until a

Table 2 Kappa coefficients for BAP anatomical and form categories

Category	Occurrence		Temporal precision (ms)			
	Proportion occurrence	κ_{occ}	Proportion duration	κ_{t1} (40 ms)	κ_{t5} (200 ms)	κ_{t11} (440 ms)
Gaze						
Straight forward	.89	.63	.80	.71	.73	.76
Upward	.12	.64	.08	.71	.72	.73
Downward	.05	.74	.01	.75	.86	.95
Averted sideways	.14	.86	.04	.75	.80	.81
Eyes closed	.12	.72	.06	.75	.78	.83
Orientation						
Head orientation	1[a]	.70	1[a]	.57	.59	.62
Trunk orientation	1[a]	.72	1[a]	.75	.77	.77
Symmetry						
Arm posture symmetry	.46	.52	.40	.47	.47	.47
Arm action symmetry	.39	.63	.23	.78	.82	.85
Arm posture						
Arm(s) at side	.39	.96	.36	.95	.96	.96
Arm(s) in front	.13	.94	.11	.92	.92	.93
Arm(s) in pocket(s)	.07	1	.06	1	1	1
Arm(s) at waist	.07	1	.06	.97	.97	.97
Head posture						
Lateral head turn	.14	.63	.10	.59	.61	.63
Lateral head tilt	.17	.90	.18	.93	.94	.94
Vertical head tilt	.15	.80	.14	.81	.82	.83
Frontal neck posture	.10	.60	.09	.60	.61	.62
Trunk posture						
Frontal trunk lean	.15	.74	.11	.72	.75	.78
Lateral trunk lean or rotation	.07	.71	.07	.80	.81	.82
Whole body posture						
Frontal body movement	.10	.68	.06	.58	.60	.61
Lateral body movement	.10	.58	.06	.47	.48	.48
Head action						
Head articulation	.21	.88	.10	.83	.87	.92
Lateral head rotation or tilt	.05	.85	.01	.72	.84	.94
Head shake vertical or lateral	.14	.97	.08	.88	.91	.96
Trunk action						
Trunk articulation	.09	.63	.05	.58	.61	.63
Forward trunk leaning	.06	.65	.02	.47	.52	.56
Arm action						
Left arm articulation	.62	.89	.47	.91	.93	.95
Right arm articulation	.59	.85	.43	.82	.85	.88
Finger(s) articulation	.53	.75	.29	.62	.68	.73
Wrist articulation	.30	.67	.14	.61	.66	.72
Elbow articulation	.54	.95	.33	.84	.88	.92

Table 2 continued

Category	Occurrence		Temporal precision (ms)			
	Proportion occurrence	κ_{occ}	Proportion duration	κ_{t1} (40 ms)	κ_{t5} (200 ms)	κ_{t11} (440 ms)
Shoulder articulation	.49	.84	.27	.76	.81	.86
Arm(s) upward	.39	.85	.14	.73	.79	.83
Arm(s) downward	.34	.82	.11	.68	.76	.83
Arm(s) forward	.30	.75	.09	.65	.73	.78
Arm(s) backward	.28	.67	.08	.57	.63	.69
Arm(s) left	.35	.81	.13	.70	.76	.82
Arm(s) right	.36	.81	.14	.71	.77	.83
Arm(s) away from the body	.37	.79	.12	.66	.74	.82
Arm(s) toward the body	.31	.75	.09	.66	.72	.77
Arm(s) vertical repetition	.13	.68	.09	.67	.69	.71
Arm(s) frontal repetition	.08	.86	.05	.84	.86	.89
Arm(s) hold	.31	.80	.12	.75	.82	.88
Arm(s) retraction	.16	.80	.04	.64	.73	.83
Other						
Touch	.24	.85	.12	.84	.88	.90
Knee bend	.06	.65	.02	.53	.58	.62
Leg movement	.40	.78	.21	.62	.68	.74

Column labels are as follows: "Proportion occurrence" = the proportion of portrayals ($N = 151$) where a behavior category was coded at least once; "κ_{occ}" = kappa computed on the occurrences of a behavior category; "Proportion duration" = the proportion of frames where a behavior category was coded ($N = 9,433$); "κ_{t1}" = kappa computed with a tolerance window of 1 frame (i.e., 40 ms); "κ_{t5}" = kappa computed with a tolerance window of 5 frames (i.e., 200 ms); "κ_{t11}" = kappa computed with a tolerance window of 11 frames (i.e., 440 ms)

[a] Orientation of the head and trunk was coded continuously (either as facing or averted), therefore the proportion of occurrence and duration equals 1

limited set of meaningful categories was reached. We refer to the Electronic Supplementary Material (Appendices A–E, available at the online journal's website or at http://www.affective-sciences.org/gemep/BAP_coding) for a detailed overview of the reliability results of all elementary and pooled variables.

Results

Below we summarize the intercoder reliability results of the behavior categories organized in sections representing behavior groups. These include occurrence agreement and temporal agreement (precision of coding). Table 2 gives a detailed overview of the results of the behavior categories situated on the anatomical and form description level. Table 3 lists the results of the behavior categories described on a functional level. In the last section, we present the frequency-based agreement of the segmentation of action units.

Table 3 Kappa coefficients for BAP function categories

Category	Occurrence agreement			Temporal precision (ms)			
	Proportion occurrence	κ_{occ} scores 3–5	κ_{occ} scores 1–5	Proportion duration	κ_{t1} (40 ms)	κ_{t5} (200 ms)	κ_{t11} (440 ms)
Emblem	.09	.90	.76	.06	.75	.79	.84
Illustrator	.18	.65	.47	.11	.35	.39	.45
Beat	.07	.65	.61	.06	.63	.68	.74
Deictic	.06	.53	.50	.02	.38	.45	.49
Manipulator	.09	.79	.56	.04	.57	.64	.74

Column labels are as follows: "Proportion occurrence" = the proportion of portrayals ($N = 151$) where a behavior category was coded at least once; "κ_{occ} scores 3–5" = kappa computed on the occurrences of a behavior category coded with a saliency between 3 and 5; "κ_{occ} scores 1–5" = kappa computed on the occurrences of a behavior category coded with a saliency between 1 and 5; "Proportion duration" = the proportion of frames where a behavior category was coded ($N = 9,433$); "κ_{t1}" = kappa computed with a tolerance window of 1 frame (i.e., 40 ms); "κ_{t5}" = kappa computed with a tolerance window of 5 frames (i.e., 200 ms); "κ_{t11}" = kappa computed with a tolerance window of 11 frames (i.e., 440 ms)

Gaze

Occurrence kappas for downward gaze, averted sideways and eyes closed were above .70 indicating excellent reliability. Only straight forward and upward gaze had lower kappas of .63 and .64 respectively. Temporal precision of coding gaze direction was already good at the zero tolerance level, and climbed further upward with increasing acceptance intervals, often reaching excellent reliability (Table 2). However, the durations of downward and averted gaze were too short (below 5% of the total duration) to allow interpretation.

Orientation and Symmetry

Intercoder agreement on occurrence was high for coding head and trunk orientation (Table 2). Temporal agreement for head orientation was .62 at the highest interval (440 ms) and dropped below .60 as the tolerance window decreased in size. Temporal agreement for trunk orientation on the other hand remained very high (above .70).

Posture and action symmetry is a quality of a posture configuration or of an action movement pattern and is thus coded conditional to the occurrence of an arm posture or action. We assessed coding agreement for symmetry only on portrayals or time segments where both coders indicated the presence of an arm posture or action. Despite high actual agreement of occurrence, arm postures were almost always symmetrical resulting in very high proportion of chance agreement, which has an undermining effect on the kappa (.52). The same restricting effect applies to the time-based kappa, which could no longer benefit from enlarging the tolerance window (flat temporal precision of .47). The proportions of arm action symmetry and asymmetry were more balanced so the kappas can be better interpreted. The occurrence coding of arm action symmetry can be considered reliable but not without concern (.63). Assessment of the temporal agreement showed very good reliability with kappas above .70.

Posture

For each arm posture category we pooled the data from the left and the right arm. One category (arm(s) behind back) did not occur often enough to interpret reliability. All other

arm posture categories were coded with an excellent to perfect agreement at both the portrayal level and the most stringent temporal level (Table 2).

The occurrence agreement was very high for the head posture categories lateral head tilt (.90) and vertical head tilt (.80), but was rather low for lateral head turn (.63) and frontal neck posture (.60). Temporal agreements were similar and did not show large improvements following increase of tolerance window.

The postural categories trunk erect, trunk bend or even the combined variable of spine and chest movement failed to reach the 5% threshold of proportion of occurrence and proportion of total duration, so we cannot evaluate reliability on the basis of this dataset. Frontal and lateral trunk lean posture was reliably coded. The agreement on postural leaning of the whole body was lower and problematic for lateral whole body posture in particular (Table 2). Leg movement was coded reliably despite limited visual information.

Action Articulation

Action articulation was reliably coded for all articulators (Table 2). Some disagreement was found for trunk action behaviors, both in the coding of occurrence as well as in temporal coding precision. On the other hand, head and arm action articulation had excellent occurrence reliability and temporal precision. Especially elbow and shoulder articulations were reliably coded even at the zero tolerance level.

Action Form

Most arm action categories at the form level had acceptable to excellent occurrence reliability (between .67 and .95, see Table 2). The time-based kappas indicating temporal precision of this group of behaviors resembled the pattern of occurrence kappas and showed considerable variation between tolerance windows. Many head and trunk action categories were too infrequently coded to compute occurrence kappa. However, categories that were coded in at least 5% of the portrayals showed excellent (above .75) occurrence reliability for head action and somewhat limited for trunk action (between .60 and .70). Similarly, temporal precision coding was excellent for head action categories but rather problematic for trunk action categories (max. .63).

Action Function

Since for each functional code the level of saliency was indicated, we computed two occurrence kappa coefficients: (a) on the whole range of behaviors, coded from 1 (*very subtle*) to 5 (*very pronounced*), and (b) on a selection of behaviors that were coded at least 3 by one or both the coders (Table 3). Comparison of the results clearly show that reliability increased for all functional categories when subtle behaviors are excluded. Further inspection of the confusions (number of coded occurrences coded by only one coder, see Appendix A of the Electronic Supplementary Material) revealed that the second coder was considerably more conservative, which either reflected more uncertainty (which led to refraining from coding), or a failure to detect a true presence (more false negatives).

The occurrence reliability varied considerably between functional categories. Emblems showed excellent reliability in both occurrence kappa computations. Manipulators had excellent reliability only when restricting to the upper range of the saliency scale, as large disagreements occurred on the coding of subtle behaviors. Also illustrators were reliably

coded when restricting to the cases with a saliency of 3 or more, although the kappa should still be taken with some concern. Similarly, beats were rather reliably coded when taking into account the whole range. Deictics failed to reach good reliability in both occurrence kappa computations.

These differences in reliability between functional categories were replicated in results on temporal agreement, although the kappas of the categories deictic and manipulator may not be stable as they represented only a small fraction of the behavior repertoire of our corpus (<5%). Increasing tolerance windows had a beneficial effect on all categories. Emblems were reliably coded already at the zero tolerance level and reached excellence at the highest tolerance level. Temporal agreement of illustrators remained very low, except for the subcategory of beats, which had a good and increasing reliability over time windows.

Segmentation

We assessed agreement on the coding of transition and configuration phases of pooled posture categories with an occurrence kappa of at least .60 and a minimal occurrence of 5%. The average occurrence proportion of agreement of these eleven[4] postures was .95 for both posture transitions and configurations. However, the kappa coefficients on the transition and configurations could not be calculated or interpreted due to highly unequal distributions. This resulted from the short durations of the portrayals in which postures often only included a configuration phase, or only or no transition phase.

We calculated agreement on the segmentation of the action units with the intercoder correlation based on the frequencies of subunits per portrayal. The Pearson correlation coefficients ($N = 151$, $df = 149$) were high for all anatomical groups (head, .85; trunk, .74; left arm, .91; right arm, .84). This indicates that there is a strong positive linear relationship between the numbers of subunits coded by both coders for all anatomical groups.

Discussion

The BAP categories listed in Table 1 represent all the behaviors that were coded using the BAP. Some coded behaviors could currently not be evaluated because they occurred with a low proportion (<5% of the entire corpus of emotional portrayals), which precluded accurate reliability estimation. We assessed intercoder reliability of occurrence and temporal precision by computing Cohen's kappa for each category. In general, the occurrence kappa coefficients mainly varied from good to perfect depending on the type of behavior (Tables 2 and 3).

Highest agreement was found for arm postures and arm and head action categories at the anatomical form level. Two form related categories, lateral whole body posture and arm posture symmetry, were not coded reliably. Improvement in coding symmetry could be made by applying an ordinal scale for gradient scoring or by providing the coding option "*not codable*". It is not entirely obvious why lateral whole body posture could not be reliably coded. We think that the complexity of whole body posture categories in terms

---

[4] These categories were: lateral head turn, lateral head tilt, vertical head tilt, frontal neck posture, frontal trunk lean, lateral trunk lean or rotation, frontal body movement, arms at side, arms in front, arms in pocket, and arms at waist.

of multiple possible articulations may have had a negative impact on intercoder agreement in combination with the limited visibility of the lower limbs from the knees down. A lack of visibility may also be responsible for considerable disagreement on the occurrence and timing of knee bend. Gaze detection was successful from the full-body view. Upward gaze was often very subtle, only slightly above the threshold, which may have led to some confusion with forward gaze and subsequently affected reliability of these two categories compared to the other gaze categories. Confusion may also have occurred in the event of co-occurring articulations of anatomically related body parts. This could have had a negative impact on the reliability (fluctuating around .60) of the coding of frontal neck posture (to separate from frontal trunk lean), lateral head turn (to separate from lateral head tilt), and knee bend (to separate from other lower limb movement). Furthermore, some behaviors occurred very rarely such as trunk actions (only forward leaning reached the 5% threshold), which could have lowered coders' attention and subsequently their agreement to this kind of behavior.

The reliability results of the functional coding showed that ambiguity of the behavior, as one might expect, had a negative effect on the reliability and seemed to play a major role for manipulators in particular. Coding on an ordinal scale or refraining from coding subtle or ambiguous movements can thus improve reliability. However, an additional factor seemed to play a role in the low intercoder agreement of illustrators and deictics in particular. Unlike emblems and manipulators, illustrators are closely linked to the accompanying verbal message, so this group of behaviors is particularly difficult to recognize when no semantic information is available about the referent, and a higher level of inference is thus needed. Inference from the speech content is especially problematic for mute coding. This is shown in our results by the difference in reliability between beats and deictics. Beats can be more easily inferred from the movement pattern because they illustrate temporal aspects of speech. Deictics are more closely tied to the speech content and the presence of objects in space, which were both absent in the GEMEP corpus. Not surprisingly, the coding of emblems did not suffer from the problem of reference because the meaning of the movement is by definition encapsulated in the movement form itself. The inherent opacity of illustrator coding in this study also explains why the frequencies of coded occurrences are generally very low given the fact that this group represents one of the largest and most frequently occurring communicative behavior types (Harrigan 2005). It reflects conscientiousness on behalf of the coders who were instructed to refrain from coding a functional category when its function could not be ascertained. In sum, successful coding of illustrator behaviors and its subtypes requires a minimum of referential information from speech or another semantic context.

Finally, given that temporal precision of action coding was reliable, we can interpret the high correlations of segment frequency as a sign of good agreement on action segmentation. Nevertheless, these results provide partial evidence of action segmentation agreement. Also the high overall percentages of agreement of posture transition and configuration is promising but should be interpreted with care given that the kappa could not be calculated or interpreted. The development of a more refined measure that compares the time-points of the segment boundaries falls out of scope for this study but should be pursued in future studies.

Methodological Benefits of the BAP for Emotion Research

From the reliability results we can conclude that, with exception of the few categories for which reliability was problematic or could not be determined, the BAP coding system can

be adopted in empirical studies where both general occurrences or timing and durations of body postures or actions are crucial to research hypotheses. As our major goal is to foster emotion expression research that includes body movement, we outline four important ways as in which the BAP coding system has been designed to reach this goal.

First, the BAP coding system is founded on a reliable observation protocol described in detail in the coding manual and guidelines. This limits the possibility of observer inference in describing emotional expressions and facilitates systematic replication and comparison of results across different studies.

Second, contrary to gesture descriptions proposed by cognitive linguists (e.g., McNeill 2005), action and posture descriptions resulting from the BAP coding system are independent of other modalities, speech in particular. Speech-independent coding was only problematic for the functional category of illustrators, except beats. Descriptions are based on information obtained from body movement alone and can thus be evaluated against other sources of information, and used to test hypotheses about the role of specific features of BAP in the encoding and decoding of emotion.

Third, the application of the BAP coding system is not restricted to the study of particular modes of expression or emotion domains selected on the basis of a priori theoretical questions. Earlier approaches often focus on one particular subsystem, such as posture (Mehrabian and Friar 1969; Tracy and Robins 2007) or action (Ekman and Friesen 1972), without specifying differential inclusion criteria. In contrast, the scope of behaviors that can be studied using the BAP coding system is large, ranging from head to leg movement including postures and actions including gestures. Movement is described in terms of locus of articulation of the major body joints. Like FACS, the BAP coding system is anchored in human anatomy and descriptions are thus closely related to the mechanisms of movement production. It includes aspects from existing systems (notably from Ekman and Friesen 1972; Frey and Pool 1976; Mehrabian 1972; Wallbott 1998), and integrates them into a single system where behavior groups (action, posture) and levels (anatomical, form, function) are distinguished. We did not pursue movement quality description here as spatio-temporal movement parameters can be reliably measured via complementary methods, (described and obtained for the GEMEP dataset in Dael and Scherer 2011; Glowinski et al. 2011).

A possible limitation is that the current version of the BAP coding system does not allow exhaustive description of all kinds of body movement one could possibly be interested in. For example, because the camera settings did not allow recording of articulation from the knees down, we could not perform exact anatomical coding of the categories of whole body posture and leg movement. The coding system neither allows refined descriptions of small articulators such as the wrist and finger joints. BAP descriptions of movement direction are not able to capture the immense complexity of possible movement trajectories and paths especially of hand action. Categorical coding systems that focus on specific hand shapes, orientations, positions and movement trajectories have been established in the fields of sign language (e.g., the "HamNoSys" system, Prillwitz 1989) and linguistics and gesture studies (Bressem 2008). The detailed and physically based descriptions of hand movement form from these systems provide complementary information to the descriptions obtained with the BAP coding system. It remains to be investigated whether specific hand movement forms are relevant cues for emotion expression. However, few studies on emotional prosody in sign language suggest that emotion can be conveyed via hand movement qualities such as shape and position in space (Hietanen et al. 2004; Reilly et al. 1992).

Fourth, the inclusion of segmentation criteria and implementation of the system in the ANVIL software platform allowed time-aligned description at a micro level. We assessed intercoder reliability on occurrence and temporal precision, and preserved only categories with a high reliability index. As such, the BAP coding system provides reliable data on occurrences, frequencies, and temporal properties of BAP such as timing and duration. This answers an important need in emotion research (Scherer and Ellgring 2007), allowing for the first time to measure the temporal unfolding and synchronization of expressive body movement.

Research Applications

The BAP coding system is designed for, but not restricted to, the study of multiple emotions. The system can be used for the purpose of exploration as well as hypothesis testing on emotion or other variables of interest. As a first application, we used the system for measuring bodily correlates of several emotions and their power in differentiating emotions (Dael et al. 2011). Analyses from this encoding study revealed several patterns of body movement, which were systematically related to emotion in the GEMEP corpus. Despite the existence of individual differences, emotion was the main determinant in behavior patterning allowing accurate discrimination of several emotions and meaningful confusions between related emotions. Few emotions were specifically related to a particular response pattern, whereas most emotions were variably expressed by multiple patterns, linking emotions such as pride and elated joy, or sadness and relief.

The coding system's built-in flexibility allows application to various material including field recordings. Only a minimum of adjustments should be made to fit the scheme to the new material (e.g., on the number of interlocutors for coding gaze and orientation). Depending on the research questions one can also choose to code behavior of selected body parts, at the functional or form level. Finally a selection can be made of only action or posture behaviors following the provided definitions. In sum, subsystems can easily be created on the basis of our validated coding system to fit the material under investigation.

The BAP coding system can be also used for dynamic multimodal analyses. If recorded on a synchronous time basis, facial and speech signals can even be visually integrated with the BAP coding on the current Anvil platform via its plug-in features.

Finally, the system's independence of speech enables application to non-linguistic settings, including behavioral research on preverbal children and non-human primates of which the motor system shows considerable overlap with that of humans. After all, coding schemes are vital tools in ethology (called ethograms) for systematically gathering animal behavior data. The human-based FACS has recently been adapted to infants (Oster 2008) and chimpanzees (Vick et al. 2007). Validation of the BAP coding system on primate behavior may provide a similar pathway for doing comparative research based on a shared methodology.

Guidelines for Using the BAP Coding System

Researchers interested in using the BAP coding system can obtain the XML coding scheme upon request to the authors. Careful training in correct application of the definitions and procedures is essential to producing replicable results.

The video material eligible for using the BAP coding system is any reasonably visible human body movement recorded in the field or in the laboratory. The body can be in any position as this does not a priori restrict the use of the BAP coding system. In the case of

seated or other positions, lower limb movement may be restricted. Furthermore, in order to obtain an accurate description of movement in all three spatial dimensions, information should be available from several camera angles. If only a frontal view is available then movement in the frontal plane should not be coded or at least handled with care. Upon minor adjustment in the XML file, the BAP coding system can be used to code social interaction with more than two interaction partners.

Anvil is currently used as an interface for using the BAP coding system because it allows for high precision time-locked coding on multiple parallel levels. Before the start of this study the Anvil software was customized[5] according to the current research demands as well as common requirements in nonverbal behavior research, including features such as a synchronized view of multiple videos and a frame-based output format, and general improvement of handling existing features. Anvil can be downloaded from http://www. anvil-software.de/. The BAP coding system can be used with other applications such as ELAN (Hellwig et al. 2010) although this has not been tested and would require conversion of the XML file.

Conclusion

In this study we introduced the BAP coding system as a tool for emotion expression research and potentially other areas of nonverbal behavior. We provided evidence of its coding reliability of occurrence, precision, and segmentation. Though the reliability results are very promising, further validation is obviously needed. Continued efforts should especially be made on the assessment segmentation reliability. By taking into account some of the major methodological concerns and desiderata in affective sciences, we hope that this new coding system will facilitate inclusion of the bodily modality in emotion research, which is principally orientated on the face and voice as expressive modalities. More emotion production studies are urgently needed so as to enhance our understanding of the mechanism of emotion expression, which can in turn inform theories of emotion perception and attribution.

References

Arieti, S. (1959). Manic-depressive psychosis. In S. Arieti (Ed.), *American handbook of psychiatry* (pp. 439–454). New York: Basic Books.

Bakeman, R., & Gottman, J. M. (1987). Applying observational methods: A systematic view. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 818–854). New York: Wiley.

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2011). Introducing the Geneva multimodal expression corpus of experimental research on emotion perception. *Emotion*. Advance online publication. doi:10.1037/a0025827.

---

[5] We are grateful to Michael Kipp for implementing the required modifications into the Anvil software package.

Bänziger, T., & Scherer, K. R. (2010). Introducing the Geneva multimodal emotion portrayal (GEMEP) corpus. In K. R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for affective computing: A sourcebook* (pp. 271–294). Oxford, UK: Oxford University Press.

Bente, G., Krämer, N. C., Petersen, A., & de Ruiter, J. P. (2001). Computer animated movement and person perception: Methodological advances in nonverbal behavior research. *Journal of Nonverbal Behavior, 25*(3), 151–166.

Birdwhistell, R. L. (1970). *Kinesics and context: Essays on body motion communication*. Conduct and communication. Philadelphia: University of Pensylvania Press.

Bressem, J. (2008). *Notating gestures—Proposal for a form based notation system of coverbal gestures.* Unpublished manuscript. Retrieved from http://www.janabressem.de/publications.html.

Bull, P. (1987). *Posture and gesture*. Oxford, UK: Pergamon Press.

Calbris, G., & Doyle, O. (1990). *The semiotics of French gestures.* Advances in semiotics. Bloomington: Indiana University Press.

Cohen, J. A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46.

Coombes, S. A., Cauraugh, J. H., & Janelle, C. M. (2007). Dissociating motivational direction and affective valence. *Psychological Science, 18*(11), 938–942.

Dael, N., Mortillaro, M., & Scherer, K. R. (2011). Emotion expression in body action and posture. *Emotion.* Advance online publication. doi:10.1037/a0025737.

Dael, N., & Scherer, K. R. (2011). Expression of emotion in gesture dynamics. Manuscript in preparation.

de Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior, 13*(4), 247–268.

Duncan, S., McNeill, D., & McCullough, K. (1995). How to transcribe the invisible-and what we see. In D. O'Connell, S. Kowal, & R. Posner (Eds.), *Zeichen für Zeit: Zur Notation und Transkription von Bewegungsabläufen* (special issue of KODIKAS/CODE., pp. 18:75–94). Tübingen: Gunter Narr Verlag.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*, 49–98.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication, 22*(4), 353–374.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior, 1*(1), 56–75.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Tomkins, S. S. (1971). Facial affect scoring technique. A first validity study. *Semiotica, 3*(1), 37–58.

Fleiss, J. L. (1981). Statistical methods for rates and proportions. In *Wiley series in probability and statistics.* New York: Wiley.

Frey, S., & Pool, J. (1976). *A new approach to the analyses of visible behavior.* Bern: Forschungsberichte aus dem Psychologischen Institut, Universität Bern.

Frey, S., & von Cranach, M. (1973). A method for the assessment of body movement variability. In I. Vine & M. von Cranach (Eds.), *Social communication and movement* (pp. 389–418). London: Academic Press.

Glowinski, D., Dael, N., Camurri, A., Volpe, G., Mortillaro, M., & Scherer, K. R. (2011). Towards a minimal representation of affective gestures. *IEEE Transactions on Affective Computing, 2*(2), 106–118.

Gross, M. M., Crane, E. A., & Fredrickson, B. L. (2010). Methodology for assessing bodily expression of emotion. *Journal of Nonverbal Behavior, 34*, 223–248.

Harrigan, J. A. (2005). Proxemics, kinesics, and gaze. In J. A. Harrigan, R. Rosenthal, & K. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 137–198). New York: Oxford University Press.

Hellwig, B., van Uytvanck, D., & Hulsbosch, M. (2010). *ELAN–linguistic annotator.* Retrieved from http://www.lat-mpi.eu/tools/elan/.

Hietanen, J. K., Leppänen, J. M., & Lehtonen, U. (2004). Perception of emotions in the hand movement quality of Finnish sign language. *Journal of Nonverbal Behavior, 28*(1), 53–64.

Hinson, M. M. (1977). *Kinesiology*. Dubuque, IA: W.C. Brown.

Hutchinson, A. (1961). *Labanotation*. New York: New Directions.

Izard, C. E. (1979). *Facial expression scoring manual (FESM)*. Newark, DE: University of Delaware.

Izard, C. E. (1983). *The maximally discriminative facial movement coding system (MAX)*. Newark, DE: Instructional Resource Center, University of Delaware.

James, W. T. (1932). A study of the expression of bodily posture. *Journal of General Psychology, 7*, 90–94.

Juslin, P. N., & Scherer, K. R. (2005). Vocal expression of affect. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65–135). Oxford, UK: Oxford University Press.

Kipp, M. (2001). Anvil—A generic annotation tool for multimodal dialogue. In *Proceedings of the 7th European conference on speech communication and technology (Eurospeech)* (pp. 1367–1370).

Kipp, M. (2003). *Anvil 4.0 annotation of video and spoken language*. Retrieved from http://www.dfki.de/~kipp/anvil.

Kipp, M. (2007). *Anvil: Annotation of video and spoken language (version 4.0)*. Retrieved from http://www.anvil-software.de/.

Kipp, M., Neff, M., & Albrecht, I. (2007). An annotation scheme for conversational gestures: How to economically capture timing and form. *Journal on Language Resources and Evaluation: Special Issue on Multimodal Corpora, 41*, 325–339.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction* (pp. 23–35). Berlin: Springer.

Laban, R. (1956). *Principles of dance and movement notation*. London: MacDonald & Evans.

Laban, R. (1975). *Principles of dance and movement notation* (Vol. 2). London: MacDonald & Evans.

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

Mehrabian, A. (1972). *Nonverbal communication*. Chicago: Aldine-Atherton.

Mehrabian, A., & Friar, J. T. (1969). Encoding of attitude by a seated communicator via posture and position cues. *Journal of Consulting and Clinical Psychology, 33*(3), 330–336.

Neumann, D. A. (2002). *Kinesiology of the musculoskeletal system: Foundations for physical rehabilitation*. Philadelphia: Mosby.

Omlor, L., & Giese, M. A. (2006). Unsupervised learning of spatio-temporal primitives of emotional gait. *Perception and Interactive Technologies, Proceedings, 4021*, 188–192.

Oster, H. (2008). *Baby FACS: Facial Action Coding System for infants and young children*. New York: New York University.

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition, 82*(2), B51–B61.

Prillwitz, S. (1989). HamNoSys. Version 2.0; Hamburg notation system for sign languages. An introductory guide. In *International studies on sign language and communication of the deaf* (Vol. 5, p. 46). Hamburg: Signum.

Reilly, J. S., McIntire, M. L., & Seago, H. (1992). Affective prosody in American sign language. *Sign Language Studies, 75*, 113–128.

Rosenfeld, H. M. (1982). Measurement of body motion and orientation. In K. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 199–286). Cambridge: Cambridge University Press.

Sayette, M. A., Cohn, J. F., Wertz, J. M., Perrott, M. A., & Parrot, D. J. (2001). A psychometric evaluation of the facial action coding system for assessing spontaneous expression. *Journal of Nonverbal Behavior, 25*(3), 167–185.

Scherer, K. R. (1979). Nonlinguistic vocal indicators of emotion and psychopathology. In C. E. Izard (Ed.), *Emotions in personality and psychopathology* (pp. 493–529). New York: Plenum.

Scherer, K. R. (1982). Methods of research on vocal communication. Paradigms and parameters. In K. R. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 136–198). Cambridge: Cambridge University Press.

Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion, 15*(2), 123–148.

Scherer, K. R., & Ellgring, H. (2007). Multimodal expression of emotion: Affect programs or componential appraisal patterns? *Emotion, 7*(1), 158–171.

Scherer, K. R., & Wallbott, H. G. (1985). Analysis of nonverbal behavior. In T. A. van Dijk (Ed.), *Handbook of discourse analysis* (pp. 199–230). London: Academic Press.

Tracy, J. L., & Robins, R. W. (2007). The prototypical pride expression: Development of a nonverbal behavior coding system. *Emotion, 7*(4), 789–801.

Vick, S. J., Waller, B. M., Parr, L. A., Smith Pasqualini, M. C., & Bard, K. A. (2007). A cross species comparison of facial morphology and movement in humans and chimpanzees using the Facial Action Coding System (FACS). *Journal of Nonverbal Behavior, 31*(1), 1–20.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28*(6), 879–896.

Wolff, C. (1951). *The hand in psychological diagnosis*. London: Methuen.