The EU-Emotion Stimulus Set: A validation study

Helen O'Reilly[1,2] · Delia Pigat[1] · Shimrit Fridenson[5] · Steve Berggren[3,4] · Shahar Tal[5] ·
Ofer Golan[5] · Sven Bölte[3,4] · Simon Baron-Cohen[1,6] · Daniel Lundqvist[3]

Abstract The EU-Emotion Stimulus Set is a newly developed collection of dynamic multimodal emotion and mental state representations. A total of 20 emotions and mental states are represented through facial expressions, vocal expressions, body gestures and contextual social scenes. This emotion set is portrayed by a multi-ethnic group of child and adult actors. Here we present the validation results, as well as participant ratings of the emotional valence, arousal and intensity of the visual stimuli from this emotion stimulus set. The EU-Emotion Stimulus Set is available for use by the scientific community and the validation data are provided as a supplement available for download.

Keywords Emotion recognition · Emotion expression · Emotion stimulus set · Validation

✉ Helen O'Reilly
h.o'reilly@ucl.ac.uk

1 Autism Research Centre, University of Cambridge, Cambridge, UK

2 Institute for Women's Health, University College London, London, UK

3 KIND, Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

4 Division of Child and Adolescent Psychiatry, Stockholm County Council, Stockholm, Sweden

5 Bar Ilan University, Ramat Gan, Israel

6 Cambridgeshire and Peterborough NHS Foundation Trust, CLASS Clinic, Cambridge, UK

A number of emotion stimulus sets have been developed for research in human emotion perception and affective computing. The majority of these sets focus on static facial expressions of basic emotions (*anger*, *disgust*, *fear*, *happiness*, *sadness* and *surprise*)—for example, the Radboud Faces Database (Langner et al., 2010), the NimStim set of facial expressions (Tottenham et al., 2009), the 3D Facial Emotional Stimuli (Gur et al., 2002), the Karolinska Directed Emotion Faces (Lundqvist, Flykt, & Öhman, 1998) and the Pictures of Facial Affect (Ekman & Friesen, 1976). However, an increasing number of dynamic emotion stimulus sets have recently been developed particularly for use in affective computing—for example, the Cohn–Kanade AU-Coded Facial Expression Database (Kanade, Cohn, & Tian, 2000; Lucey et al., 2010) and the Video Database of Moving Faces & People (O'Toole et al., 2005). See Bänziger, Mortillaro, & Scherer (2012) for an overview of emotion stimulus sets.

The stimulus sets mentioned above predominantly focus on *facial* expressions of emotion, a fact that reflects the greater attention of emotion research to facial expressions than to other expressive modalities, such as body gestures. Faces are a salient signal of another's affective state, and the ability to perceive and understand facial expressions is an important skill for social interaction. Furthermore, certain clinical populations demonstrate impairments in the recognition of facial expressions (Bölte & Poustka, 2003; Comparelli et al., 2013; Meletti et al., 2003; Sucksmith, Allison, Baron-Cohen, Chakrabarti, & Hoekstra, 2013). However, previous research has highlighted that emotion recognition is influenced by an interaction between the expressive modalities of face, voice and body (Meeren, van Heijnsbergen, & de Gelder, 2005; Regenbogen et al., 2012; Van den Stock, Righart, & de Gelder, 2007). Hence, an emotion stimulus set that permits the investigation of emotion perception across the expressive modalities of face, voice and body, both individually and

combined, would be a valuable resource for the scientific community. In addition, emotion expression is dynamic in nature, and previous research has suggested that the use of dynamic emotion stimuli is more ecologically valid, results in better emotion recognition (Ambadar, Schooler, & Cohn, 2005; Trautmann, Fehr, & Herrmann, 2009; Weyers, Mühlberger, Hefele, & Pauli, 2006), and also activates a wider neural network compared to static emotion stimuli (Kilts, Egan, Gideon, Ely, & Hoffman, 2003; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Trautmann et al., 2009).

The EU-Emotion Stimulus Set was created as part of the ASC-Inclusion project within the European Community's Seventh Framework Programme (FP7/2007-2013; www.asc-inclusion.eu). The aim of the ASC-Inclusion project was to create an online socio-emotional training tool for children with a diagnosis of autism spectrum condition (ASC). In the course of the development of this online training tool, an emotion stimulus set was required that portrayed a range of emotions and mental states through the three expression modalities of face, voice and body gesture, as well as contextual social scenes. However, no such stimulus set was available covering both the range of emotions/mental states and modalities required. The only other, somewhat similar dynamic stimulus sets known to the authors that are freely available for scientific use are the Amsterdam Dynamic Facial Expression Set (Van der Schalk, Hawk, Fischer, & Doosje, 2011), the CAM Face–Voice Battery (Golan, Baron-Cohen, & Hill, 2006), and the Geneva Multimodal Emotion Portrayals Core Set (GEMEP-CS; Bänziger et al., 2012).

The EU-Emotion Stimulus Set expands on the previous dynamic emotion sets in (1) the numbers of emotions and mental states represented, (2) the age range of the actors expressing the emotions, and (3) the expression modalities through which these emotions/mental states are portrayed. In this report, we outline the development of the visual stimuli from the EU-Emotion Stimulus Set, along with the validation results and emotional ratings from typically developed adults. The EU-Emotion Stimulus Set is freely available to investigators for use in scientific research, and can be downloaded from www.autismresearchcentre.com/arc_tests.

Method

Stimulus creation

Stimulus set

The EU-Emotion Stimulus Set contains $N = 418$ visual stimuli (video clips, durations 2–52 s) of 20 different emotions and mental states, plus neutral. This stimulus set is dynamically portrayed by 19 actors through facial expressions ($n = 249$),

body gesture scenes ($n = 82$), and contextual social scenes ($n = 87$). See Appendixes A, B and C for example visual stimuli. A total of $N = 2,364$ vocal stimuli (English $n = 698$, Swedish $n = 1,012$, Hebrew $n = 654$) have also been produced and validated (the methodology and results for these vocal stimuli are currently in preparation for publication separately).

Emotions and mental states

The set consists of the following 20 emotions/mental states, plus a *neutral* state: *afraid*, *angry*, *ashamed*, *bored*, *disappointed*, *disgusted*, *excited*, *frustrated*, *happy*, *hurt*, *interested*, *jealous*, *joking*, *kind*, *proud*, *sad*, *sneaky*, *surprised*, *unfriendly*, and *worried*. These states were selected from an initially evaluated set of 27 emotions/mental states (see Lundqvist et al., 2014). ASC clinical experts ($n = 47$) and the parents of children with ASC ($n = 88$) rated these 20 emotions/mental states as being the most important for social interactions, out of the potential 27. Limited research has investigated the visual or auditory distinctiveness/uniqueness of these more complex emotions/mental states. Two prior studies that have examined more complex emotions/mental states (i.e., other than the six basic emotions: *anger*, *disgust*, *fear*, *happiness*, *sadness*, and *surprise*) showed the complex emotions/mental states investigated to be discretely identifiable through facial and vocal expression (Bänziger et al., 2012; Golan et al., 2006). In the present study, no prior assumptions were made as to whether each of these 20 emotions/mental states hold discretely identifiable facial, vocal, and body expressions, but rather, the study provided the opportunity to examine the uniqueness of these emotions/mental states across modalities while concurrently developing a valid set of emotion stimuli.

Actors

The stimuli were depicted by 19 actors of different ethnicities, age range 10–70 years old (ten female and nine male). See Table 1 for the actor demographics. The actors were recruited from professional acting agencies or drama schools within the United Kingdom. A director was hired with experience in theatre and film production to manage the filming. Approximately 80 actors auditioned for the roles. One member of the research team and the director selected the 19 actors on the basis of the quality of their audition performances. The actors were advised/guided through their performance by the director.

Instructions

The actors were first filmed performing the facial expressions, followed by the body gesture scenes, and finally the contextual social scenes. The actors performed each scene three

Table 1 Actor demographics and modalities completed

Actor	Gender	Age	Ethnicity	Facial Expressions	Body Gestures	Social Scenes
A	Female	19	Mixed Afro-Caribbean/Asian British	X	✓	✓
B	Female	37	White British	✓	✓	✓
C	Male	31	White British	✓	✓	✓
D	Female	27	Black British	✓	✓	✓
E	Female	70	White British	✓	✓	✓
F	Female	10	Black British	✓	✓	✓
G	Female	15	White British	✓	✓	✓
H	Male	62	White British	✓	✓	✓
I	Male	11	White British	✓	X	✓
J	Male	12	Mixed White/Asian British	✓	X	✓
K	Male	30	White British	✓	X	✓
L	Male	12	White British	✓	X	✓
M	Male	37	White British	✓	X	✓
N	Female	42	White British	✓	X	✓
O	Female	21	White British	✓	X	✓
P	Male	12	White British	✓	X	✓
Q	Female	10	Mixed White/Asian British	X	X	✓
R	Female	11	White British	✓	X	✓
S	Male	11	Mixed Mediterranean/Asian British	✓	X	✓

times for all modalities. No vocalizations were produced while the actors performed the facial expressions, body gesture scenes or the contextual social scenes. Filming was shot using an infinite white background.

For the facial expressions, each actor ($N = 17$) was asked to portray a subset of ten emotions/mental states, plus *neutral*. Ten of the emotions/mental states were assigned to each actor rather than the full 20, due to filming time constraints. The 20 emotions/mental states were divided into two sets of ten emotions/mental states each (three basic emotions and seven complex emotions/mental states). The actors received facial expression scripts for either Set A or B. The two sets were distributed among the actors to try to ensure even representation across genders and ages. For the facial expression scenes, a straight camera angle and standard zoom was applied, giving a frontal view of each actor with only the shoulders and head visible within the shot. Scripts that described a possible scenario in which an emotion/mental state would occur were provided to help the actors portray the different emotions/mental states. An example follows of a script for the facial expression *ashamed*: "Face: Your mum caught you eating the cookie you stole from the kitchen and she is angry." The use of scripts helped balance the intensity at which the emotions/mental states were displayed across actors, and specific instructions were also provided on intensity. The six basic emotions (*anger, disgust, fear, happiness, sadness,* and *surprise*) were each portrayed twice through facial expressions, once at a high intensity and once at a low intensity,

and the other 14 emotions/mental states were all portrayed through facial expression at a high intensity. The following example instruction was given to help guide the intensity of expression across all modalities: "High Intensity—In this situation, you are *quite* ashamed; not *a little* ashamed, not *very* ashamed, but *quite and unmistakably* ashamed."

The body gesture scenes were completed by eight of the actors. Each emotion was individually portrayed through a body gesture scene by one actor alone, with the exception of the following six emotions/mental states—*ashamed, hurt, jealous, joking, kind,* and *unfriendly*—in which two actors acted out the body gesture scene. The six emotions/mental states listed above were presented in a scene by two actors, because it was felt that these emotions were difficult to act out/interpret without the involvement of a second character to whom the emotion/mental state could be directed. These emotions/mental states depend on a social interaction for their expression. Again, the emotions/mental states were divided into four scripts, two for the individual body gesture scenes and two for the dual-actor body gesture scenes. Each actor received two scripts (a total of ten emotions/mental states plus *neutral* to portray), one for the individual body gestures and the other for the dual-actor body gesture scenes. Scripts were provided to guide the actors' performances and intensities of expression. All emotions/mental states were performed at a high intensity, and the entire body was visible within the shot.

For the contextual social scenes, one to three actors were grouped together and requested to perform a social scenario.

All 19 actors participated in the filming of the 87 contextual social scenes. Each contextual social scene depicted one to four emotions. Again, scripts were provided detailing the social interactions required. The social scenes were filmed using long to medium shots of the whole body and facial close-ups. The shots were edited together to depict the social scene. The actors were requested to act out the emotion in the scenario as naturally as possible. An example social scene script follows: "*Sneaky and Afraid*—James sneaks into room, hides a fake spider under Laura's book sneakily. Leaves room. Laura walks into the room, picks up book from desk and discovers a spider hiding underneath—she gets a fright and drops book." The actors were requested to perform all contextual social scenes at a high intensity.

Stimulus validation

Due to the vast quantity of footage produced, it was not possible to validate every emotion/mental state performance recorded. Three members of the research team individually went through the filmed footage and selected the best clip of each emotion/mental state portrayal (each emotion/mental state was performed three times by each actor) to go forward for validation. For the majority of the footage, the three raters agreed on the best emotional portrayal. For any disagreement, the item with the majority agreement was chosen or two takes of the same item were put forward for validation.

Stimulus set and surveys

The face, body gesture and contextual social scene stimuli were divided into 14 separate online surveys (six face surveys, two body gesture surveys and six social scene surveys). The stimuli for each modality were divided into the corresponding surveys, ensuring that the emotions/mental states were evenly distributed. Each of the facial expression surveys included approximately 41 stimulus items, and each of the contextual social scene/body gesture surveys included 15–30 stimulus items. Each survey took 20–30 min to complete. The 14 surveys were first developed in English and then translated into Swedish and Hebrew (using back translation) by two native speakers for each language who were also fluent in English. The 14 surveys were distributed in the UK, Sweden, and Israel separately (14 surveys × 3 languages).

Participants

Altogether, a total of 1,231 complete responses were recorded across the 14 validation surveys (803 female, 428 male) from the three data collection sites. A minimum of 54 participants completed each survey (approximately 18 participants per data collection site and survey). The average age of the participants was 44 years (SD = 16.7, range: 16–84). Participants

were recruited using existing research participant databases and university mailing lists, as well as through online resources—for example, on social media sites such as the project Facebook page.

Procedure

To estimate the validity of the EU-Emotion stimuli, we investigated whether each stimulus could be recognized as the intended emotional/mental state expression and the degree to which a stimulus conveyed an emotional impression.

Recognition task The recognition rates were investigated by using a forced choice task. For each stimulus, a video of the emotional stimuli was shown, together with six counterbalanced response options. The participants were instructed to "Please select the label which best describes what this person is expressing." The six response options consisted of the target emotion/mental state (i.e., the emotion/mental state that the actor intended to express), four control emotions/mental states, and a *None of the above* response option. The *None of the above* option was included to prevent artifactual agreement (Frank & Stennett, 2001).

In the related literature, when researchers have validated stimuli across a smaller range of emotions, such as the six or seven basic emotions (see, e.g., Goeleven, De Raedt, Leyman, & Verschuere, 2008; Langner et al., 2010), they have adopted adopt an "all against all" approach when selecting response options. For example, when presenting a *happy* stimulus, *afraid*, *angry*, *disgusted*, *happy*, *neutral*, *sad*, and *surprised* are used as response options, and the same set can be used when presenting a *neutral* or a *sad* stimulus, and so forth. Because of the large number of emotions in the EU-Emotion Stimulus Set, it was not possible to use this approach, since a 20-option input was too arduous and overwhelming for participants. To accomplish a limited set of response options per stimulus while also maintaining a fair comparison across emotions, we adopted an innovative approach.

Because the task itself consisted of selecting the label that matched with the stimulus from among a handful of response options, we wanted to present response options for each emotion/mental state that were equal in difficulty across all target emotions/mental states. To accomplish this, we used the data of another study from our group (Lundqvist et al., 2014). In this study, over 700 participants rated the similarity/dissimilarity of each of the 20 emotions/mental states involved here against all of the other 20 emotions/mental states. Since this study resulted in a 20×20 emotion similarity/dissimilarity matrix, we adopted the strategy of first specifying different ranges of similarity (corresponding to *very similar*, *quite similar*, *quite dissimilar*, and *very dissimilar*), and then (2) selected appropriate response alternatives from the similarity/dissimilarity matrix within these four ranges for

each target emotion. This selection logic was applied in order to guarantee that the sets of response options presented for each target emotion/mental state were equally easy or difficult across all emotions/mental states and stimuli (see the "Emotions Matrix" in the supplementary data for a list of the target and control emotions/mental states, along with the similarity ranges used for selection).

Emotional impression The emotional impression of each stimulus was assessed through subjective ratings of valence, arousal, and intensity. For valence, the question "How Positive or Negative is this emotional expression?" was used. The participants were instructed to answer on a scale from 1 to 5, where 1 = *Very negative* and 5 = *Very positive*. For arousal, the question "How strongly does this emotional expression make you feel?" was used. The participants were instructed to answer on a scale from 1 to 5, where 1 = *Not at all* and 5 = *Very strongly*. For intensity, the question "How intense is this emotional expression?" was used, and participants were instructed to answer on a scale from 1 to 5, where 1 = *Calm*, and 5 = *High intensity*.

Data treatment and analysis

For the recognition data, raw recognition rates were first calculated for all stimuli and all response options. Since we were using six response options, the raw scores were then adjusted for chance rates by using Cohen's kappa [True Correct = (Proportion of Raw Correct – (1/6) / (5/6)] (Tottenham et al., 2009). The average recognition scores and emotional rating scores were calculated separately per stimulus and then accumulated over emotions, separately per emotion intensity and per modality (facial expression, body expression, or social scenario). In the tables (here and in the supplementary data), true scores below 0 have been adjusted to 0. Also, to give an overview as to what degree the main dependent measures were correlated, we calculated intercorrelations for the (chance-corrected) recognition scores and the valence, arousal, and intensity ratings using Pearson's correlation coefficient. A comparison of the emotion recognition scores cross-culturally was not possible, due to variation in the age ranges and genders of participants between the data collection sites.

Results

Table 2 shows an overview of the recognition scores and emotional rating scores for each expression modality, each emotion/mental state, and (when applicable) the two intensity levels. Corresponding data are also made available on an individual stimulus item level as supplementary data (see the "Validation Data" in the supplementary data).

On the modality level, the mean chance-corrected recognition (CCR) scores for expressions made via the face were 63 % (SD = 16 %), for body gestures were 77 % (SD = 11 %), and for social scenarios were 72 % (SD = 17 %). For all modalities, most emotions/mental states resulted in CCR scores between 60 % and 90 % (see Table 2 and the supplementary data for details). However, some emotions/mental states even at an average level resulted in very low scores, as for facial expressions: *kind* (M = 9 %), *jealous* (M = 14 %), and *unfriendly* (M = 9 %) and for body gestures *jealous* (M = 3 %). These lows are the exception; a number of chance-corrected scores as high as 90 % or above were also be found, such as for facial expression *joking* (M = 90 %), body gestures *disappointed* (M = 97 %), *disgusted* (M = 90 %), *frustrated* (M = 98 %), and *kind* (M = 91 %), and social scenarios *afraid* (M = 94 %) and *frustrated* (M = 96 %). The intensity of the basic emotions also clearly influenced the results, resulting in an average of 78 % for high-intensity facial expressions and 63 % for low-intensity expressions (see Table 2 for details). The modality through which emotions/mental states were expressed also appeared to play an important role in recognition. *Unfriendly* and *kind* were both recognized poorly (9 % recognition) when expressed through facial expression. Recognition for *unfriendly* increased to 68 % when it was portrayed through body gestures, and to 62 % for social scenarios. Similarly for *kind,* recognition increased to 91 % for body gestures and to 61 % for social scenarios. *Jealous* also achieved low recognition scores through facial expressions (13 %) and body gestures (3 %), but increased to 44 % when represented within a social context. The emotion *surprised* was recognized best when it was expressed through the face (79 %) and body (75 %), but recognition decreased to 49 % when it was represented through social scenarios. Overall, emotions/mental states expressed through body gesture scenes showed the highest mean recognition scores (see "Emotion Recognition by Modality" in supplementary data).

The results also showed that errors were made across response options in a manner that was in line with expectations from how conceptually similar the control emotions/mental states were to the target emotion/mental state. Overall, the CCR scores across modalities were 4 % for Control Emotion 1 (the emotion most similar to the target emotion), 1.9 % for Control Emotion 2, 1.1 % for Control Emotion 3, and finally 0.2 % for Control Emotion 4 (the least similar to the target emotion; see the "Emotions Matrix" in the supplementary data for further details). This demonstrates that control emotion/mental state CCR scores declined with decreasing control-emotion-to-target similarity.

The correlation analyses showed a low level of intercorrelation between the recognition, valence, arousal, and intensity measures (R^2 scores on average around .05), with the exception of the correlation between arousal and intensity ratings. For these measures, the r scores were .90 on average, and between .78 and .96 over the different item categories (see Table 3).

Table 2 Summary of the validation data for each of the 20 emotions plus neutral, separately per expression modality and (where applicable) expression intensity (denoted [*]High and [**]Low)

Stimulus Information			Underlying Data		Recognition Scores		Emotion Scores					
Expression Modality	Expressed Emotion	Emotion Intensity	Number of Stimulus Items	Number of Responses per Item	Chance-Corrected Recognition Scores		Valence Ratings		Intensity Ratings		Arousal Ratings	
					M	Std	M	Std	M	Std	M	Std
Face	*All*	*All*	*SUM: 249*	*SUM: 3,220*	*63 %*	*16 %*	*2.78*	*0.19*	*3.44*	*0.24*	*3.22*	*0.21*
Face	*All*	*High*	*SUM: 200*	*SUM: 2,474*	*63 %*	*16 %*	*2.80*	*0.19*	*3.50*	*0.23*	*3.27*	*0.20*
Face	*Selection[*]*	*High*	*SUM: 51*	*SUM: 717*	*78 %*	*16 %*	*2.64*	*0.18*	*3.67*	*0.26*	*3.40*	*0.23*
Face	*Selection[**]*	*Low*	*SUM: 49*	*SUM: 748*	*63 %*	*18 %*	*2.69*	*0.17*	*3.25*	*0.28*	*3.08*	*0.25*
Face	Afraid[*]	High	8	106	78 %	13 %	2.14	0.18	3.57	0.26	3.32	0.22
Face	Afraid[**]	Low	9	141	70 %	8 %	2.3	0.1	3.33	0.17	3.16	0.17
Face	Angry[*]	High	9	101	69 %	14 %	2.02	0.16	3.69	0.33	3.34	0.21
Face	Angry[**]	Low	8	136	37 %	25 %	2.16	0.12	3.14	0.28	2.89	0.22
Face	Ashamed	High	8	104	59 %	26 %	2.17	0.21	3.54	0.18	3.35	0.14
Face	Bored	High	8	112	68 %	13 %	2.24	0.08	3.35	0.11	3.06	0.09
Face	Disappointed	High	10	126	55 %	13 %	2.18	0.15	3.46	0.22	3.22	0.19
Face	Disgusted[*]	High	8	97	82 %	23 %	1.91	0.13	3.97	0.2	3.65	0.16
Face	Disgusted[**]	Low	10	125	57 %	27 %	2.19	0.16	3.4	0.41	3.13	0.34
Face	Excited	High	9	109	79 %	14 %	4.4	0.21	4.2	0.19	3.84	0.21
Face	Frustrated	High	11	168	75 %	12 %	2.15	0.14	3.69	0.19	3.34	0.17
Face	Happy[*]	High	8	194	82 %	13 %	4.44	0.17	3.94	0.19	3.75	0.18
Face	Happy[**]	Low	7	92	59 %	20 %	3.98	0.21	3.22	0.26	3.23	0.28
Face	Hurt	High	10	111	59 %	16 %	2.12	0.19	3.45	0.24	3.36	0.16
Face	Interested	High	11	124	65 %	18 %	3.42	0.33	3.04	0.35	2.91	0.3
Face	Jealous	High	7	111	14 %	17 %	2.12	0.25	3.36	0.2	3.2	0.16
Face	Joking	High	9	120	90 %	9 %	4.27	0.13	4.08	0.12	3.72	0.14
Face	Kind	High	9	108	9 %	11 %	3.83	0.32	3.31	0.31	3.2	0.24
Face	Neutral	High	17	113	81 %	9 %	2.92	0.12	2.23	0.12	2.16	0.15
Face	Proud	High	11	123	58 %	29 %	3.95	0.29	3.67	0.28	3.39	0.3
Face	Sad[*]	High	8	95	80 %	10 %	2.05	0.1	3.46	0.22	3.3	0.23
Face	Sad[**]	Low	7	129	70 %	19 %	2.2	0.13	3.24	0.28	3.14	0.24
Face	Sneaky	High	11	111	59 %	19 %	2.93	0.21	3.18	0.19	3.00	0.17
Face	Surprised[*]	High	10	124	79 %	23 %	3.27	0.36	3.4	0.38	3.02	0.35
Face	Surprised[**]	Low	8	127	82 %	11 %	3.28	0.29	3.16	0.29	2.92	0.25
Face	Unfriendly	High	10	124	9 %	13 %	2.19	0.22	3.31	0.32	3.18	0.28
Face	Worried	High	8	95	63 %	15 %	2.16	0.11	3.52	0.25	3.27	0.24
Body	*All*	*High*	*SUM: 82*	*SUM: 1,766*	*77 %*	*11 %*	*2.99*	*0.25*	*3.39*	*0.23*	*3.01*	*0.21*
Body	*Selection[*]*	*High*	*SUM: 24*	*SUM: 549*	*80 %*	*12 %*	*2.88*	*0.30*	*3.51*	*0.31*	*3.00*	*0.27*
Body	Afraid[*]	High	5	88	82 %	18 %	2.25	0.08	3.39	0.33	2.91	0.21
Body	Angry[*]	High	3	72	78 %	7 %	2.41	0.47	3.43	0.21	3.06	0.25
Body	Ashamed	High	4	65	76 %	7 %	2.36	0.14	3.45	0.29	3.2	0.22
Body	Bored	High	4	93	86 %	7 %	2.21	0.11	3.11	0.21	2.68	0.19
Body	Disappointed	High	3	103	97 %	2 %	2.27	0.1	3.45	0.08	2.79	0.18
Body	Disgusted[*]	High	6	94	90 %	4 %	2.15	0.15	3.56	0.38	2.93	0.32
Body	Excited	High	5	88	64 %	20 %	4.22	0.18	3.84	0.3	3.21	0.27
Body	Frustrated	High	4	122	98 %	1 %	2.70	1.09	3.72	0.32	3.02	0.29
Body	Happy[*]	High	5	99	88 %	9 %	4.64	0.13	4.07	0.27	3.47	0.18

Table 2 (continued)

Stimulus Information			Underlying Data		Recognition Scores		Emotion Scores					
Expression Modality	Expressed Emotion	Emotion Intensity	Number of Stimulus Items	Number of Responses per Item	Chance-Corrected Recognition Scores		Valence Ratings		Intensity Ratings		Arousal Ratings	
					M	Std	M	Std	M	Std	M	Std
Body	Hurt	High	3	70	76 %	9 %	2.21	0.12	3.38	0.36	3.32	0.12
Body	Interested	High	3	85	70 %	8 %	3.72	0.31	3.04	0.32	2.7	0.12
Body	Jealous	High	4	62	3 %	7 %	2.39	0.09	3.22	0.26	2.94	0.21
Body	Joking	High	4	62	89 %	4 %	4.46	0.09	3.79	0.18	3.45	0.2
Body	Kind	High	4	60	91 %	3 %	4.38	0.08	3.66	0.1	3.68	0.14
Body	Neutral	High	6	94	84 %	5 %	2.93	0.04	1.88	0.13	1.81	0.13
Body	Proud	High	4	80	73 %	28 %	4.4	0.17	3.78	0.08	3.39	0.44
Body	Sad*	High	3	102	67 %	13 %	2.18	0.07	2.99	0.29	2.57	0.27
Body	Sneaky	High	3	102	79 %	6 %	2.73	0.3	3.15	0.03	2.75	0.08
Body	Surprised*	High	2	94	75 %	24 %	3.62	0.88	3.62	0.39	3.04	0.41
Body	Unfriendly	High	4	66	68 %	14 %	2.06	0.17	3.19	0.16	3.21	0.17
Body	Worried	High	3	85	77 %	27 %	2.48	0.5	3.42	0.06	3.05	0.06
Social	*All*	*High*	*SUM: 153*	*SUM: 1,282*	*72 %*	*17 %*	*3.00*	*0.25*	*3.65*	*0.25*	*3.32*	*0.23*
Social	*Selection**	*High*	*SUM: 63*	*SUM: 382*	*72 %*	*20 %*	*2.91*	*0.27*	*3.76*	*0.26*	*3.38*	*0.24*
Social	Afraid*	High	6	63	94 %	5 %	2.31	0.19	4.03	0.11	3.49	0.12
Social	Angry*	High	13	63	72 %	13 %	2.26	0.12	3.79	0.29	3.41	0.29
Social	Ashamed	High	7	62	75 %	14 %	2.31	0.17	3.54	0.22	3.28	0.2
Social	Bored	High	6	63	59 %	29 %	2.53	0.18	3.35	0.3	3.11	0.3
Social	Disappointed	High	7	61	66 %	13 %	2.55	0.25	3.41	0.43	3.1	0.43
Social	Disgusted*	High	6	64	82 %	10 %	2.22	0.25	3.71	0.13	3.29	0.13
Social	Excited	High	2	65	89 %	8 %	4.36	0.23	4.19	0.56	3.64	0.46
Social	Frustrated	High	5	63	96 %	4 %	2.24	0.14	3.94	0.25	3.42	0.2
Social	Happy*	High	22	65	66 %	25 %	4.2	0.28	3.72	0.38	3.35	0.31
Social	Hurt	High	9	65	81 %	14 %	2.23	0.18	3.6	0.17	3.47	0.18
Social	Interested	High	11	68	72 %	13 %	3.80	0.22	3.45	0.15	3.05	0.18
Social	Jealous	High	4	66	44 %	31 %	2.49	0.25	3.41	0.18	3.1	0.1
Social	Joking	High	5	63	87 %	8 %	4.13	0.15	3.85	0.14	3.4	0.14
Social	Kind	High	10	64	61 %	28 %	4.17	0.24	3.5	0.18	3.44	0.17
Social	Proud	High	3	64	70 %	12 %	3.96	0.44	3.76	0.36	3.33	0.25
Social	Sad*	High	8	65	67 %	31 %	2.4	0.26	3.28	0.43	3.21	0.31
Social	Sneaky	High	8	69	78 %	8 %	3.04	0.32	3.47	0.13	3.11	0.19
Social	Surprised*	High	8	62	49 %	33 %	4.06	0.53	4	0.19	3.51	0.26
Social	Unfriendly	High	8	62	62 %	16 %	2.19	0.16	3.39	0.17	3.36	0.21
Social	Worried	High	5	64	77 %	20 %	2.63	0.5	3.68	0.23	3.3	0.16

The table indicates the number of stimulus items included in each average and the number of participants involved, and also gives average and standard deviation scores for the recognition (CCR), valence, intensity and arousal measures. These data are also available on a stimulus item level as supplementary data. The underlined and double underlined represent the averages for that modality varied by intensity or selection of emotions

Discussion

The EU-Emotion Stimulus Set is a validated collection of 418 dynamic multimodal emotion and mental state representations displayed through facial expressions, body gestures, and contextual social scenes. Overall, moderate to high recognition scores were achieved across each of the modalities: facial expression ($M = 63$ %, $SD = 16$ %), body gesture scenes ($M = 77$ %, $SD = 11$ %), and contextual social scenes ($M = 72$ %, $SD = 17$ %). The validation results varied for each emotion/

Table 3 Summary of intercorrelations between (chance-corrected) recognition scores and valence, arousal, and intensity ratings, using Pearson's correlation coefficients

	Items	Recognition	Valence	Intensity	Arousal
Recognition	ALL		.08	.23***	.13**
	Face		.16*	.22**	.16*
	Body		.05	.15	.04
	Social		–.15	.33***	.29***
Valence	ALL	.08		.23***	.22***
	Face	.16*		.17**	.19**
	Body	.05		.37**	.42***
	Social	–.15		.17*	.12
Intensity	ALL	.23***	.22***		.90***
	Face	.22**	.17**		.96***
	Body	.15	.37**		.87***
	Social	.33***	.17*		.78***
Arousal	ALL	.13**	.22***	.90***	
	Face	.16*	.19**	.96***	
	Body	.04	.42***	.87***	
	Social	.29***	.12	.78***	

The table gives correlation scores and *p* values for significant relationships between the variables. * *p* < .05, ** *p* < .01, *** *p* < .001

mental state, depending on the expression modality through which that emotion/mental state was portrayed. For the facial stimuli, *joking* had the highest recognition rate, at 90 %. *Kind* and *unfriendly* facial expressions had the lowest recognition rates, both at 9 %. *Disappointed* and *frustrated* had the highest body gesture recognition scores, at 97 % and 98 %, respectively, whereas *jealous* portrayed through body gesture had a recognition rate of only 3 %. For the contextual social scenes, *frustrated* was the most identifiable (96 %), and *jealous* again had the lowest recognition score (44 %). The validation results provide valuable data on the uniqueness of the expressions and the modalities through which emotions/mental states are most clearly displayed. For example, *unfriendly* was not easily identified through facial expression (9 % recognition rate); however, recognition increased to 68 % when it was portrayed through body gestures. The results showed that *jealous* was the most difficult to identify, possibly due to its complex nature and need for contextual cues.

Turning to the correlation results, although a number of significant findings are reported, the effect sizes of the majority of these are small. The strongest correlation was found between the intensity and arousal scores (with emotions/mental states that scored high in arousal also scoring high in intensity across facial expressions, body gestures, and social scenes), demonstrating a strong relationship between these two emotion dimensions. Bänziger et al. (2012) also reported that the intensity ratings of their emotion stimuli were

influenced by arousal level, with low arousal being associated with low intensity.

The recognition results of the EU-Emotion Stimulus Set are slightly higher than those of Bänziger et al. (2012), whose stimulus set (GEMEP-CS) is the closest to the EU-Emotion Stimulus Set in terms of the range of emotions/mental states and modalities portrayed. Bänziger et al. validated 154 dynamic facial expression stimuli (face and upper torso visible) representing 17 different emotions, producing a mean overall uncorrected recognition score of 47 %. The difference in overall facial recognition scores between the EU-Emotion Stimulus Set (63 %) and GEMEP-CS (47 %; Bänziger et al., 2012) is likely explained by variation in the validation methodologies used between the studies, which we discuss further below.

Tottenham et al. (2009) reported an overall 79 % CCR rate for the NimStim facial stimulus set. This set consists of 672 static facial expression stimuli representing eight different emotions (the six basic emotions—*anger*, *disgust*, *fear*, *happiness*, *sadness*, *surprise*—plus *neutral* and *calm*). Since the EU-Emotion set has a wider range of emotional/mental state expressions, and at two levels of intensity for basic emotions, the comparatively lower overall recognition rate reported in our data is likely due to the greater complexity and variety of our stimuli. Indeed, the overall mean recognition rate of our *basic* facial emotion stimuli (high-intensity only) plus *neutral* was 78 %, which is comparable to the results of Tottenham et al. Similar recognition rates have also been reported by Langner et al. (2010), reporting an uncorrected mean recognition rate of 82 %, for a set consisting of the six *basic* emotions plus *neutral* and *contempt*. Finally, the validation results for the KDEF set reported by Goeleven, Raedt, Leyman, and Verschuere (2008) revealed an overall uncorrected emotion recognition rate (for the six basic emotions + *neutral*) of 72 %.

In the present study, a subset of four emotions/mental states were selected from the 20 emotion/mental state labels to act as the control response options, in addition to *None of the above*, in the forced choice validation task. The four control response options varied depending on the target emotion/mental state portrayed. The use of a forced choice task is a common method that has been used in previous emotion validation studies (Bänziger et al., 2012; Tottenham et al., 2009). The NimStim face stimulus set portrays the six basic emotions plus *neutral* and *calm*. Participants had to choose between these eight labels, plus a *None of the above* option, within this forced choice validation study (Tottenham et al., 2009). The mean corrected recognition rates were high for the individual emotions (.54–.95). This high recognition rate could possibly have been due to the limited number of response options to choose from and the discreteness of the basic emotions. Bänziger et al. had participants choose between 17 emotion labels and an *Other emotion* option. This resulted in lower mean uncorrected

recognition rates (.28–.79) than those reported by Tottenham et al. This was likely due to the fact that the emotions investigated by Bänziger et al. were greater in number, more complex (emotions beyond the six basic emotions were investigated, such as *tenderness*, *pleasure*, and *despair*) and potentially overlapped with one another (e.g., *sadness* and *despair*, *elated joy* and *amusement*, *contempt* and *irritation*), making recognition more challenging in their validation task.

When designing the validation task for the EU-Emotion set, we decided not to present the participants with the full list of 20 emotions/mental states, plus *neutral* and a *None of the above* option, following the validation design of Bänziger et al. (2012). Presenting all 20 options would be visually overwhelming and time burdening, and would potentially result in the participants not considering each of the emotion/mental state labels in turn with regard to the stimulus presented. Also it could increase the risk of participants relying on the basic emotion labels (*anger*, *disgust*, *fear*, *happiness*, *sadness*, and *surprise*) as broad categories when viewing the stimuli, rather than considering the particular characteristics of the more complex emotions/mental states. This led to the decision when selecting the control options from the similarity matrix not to include control emotions/mental states with too high a similarity rating with the target emotion/mental state. For example, for the target emotion *disappointed*, the emotion *sad* was not used as a control response option, due to its strong similarity and overlap with *disappointed* in both meaning and expression. Without sufficient context it would be difficult to distinguish these two emotions from one another, which would possibly lead to participants focusing on the basic emotion labels. This difference in methodology and also in the numbers of emotions/mental states to choose between (6 vs. 17) in the present study and that of Bänziger et al. likely explains the variation in mean recognition rates reported between these studies. In the EU-Emotion Stimulus Set validation task, the four control emotions/mental states varied for each target emotion/mental state. This raises the question of whether the recognition scores for each emotion/mental state can be compared. Although the control options did vary, their similarity with the target emotion/mental state was consistent across all target emotions/mental states, as was the ease/difficulty level of the forced choice task. This method provided a standardized measure against which each stimulus was judged and prevented some potential issues from occurring, as outlined above.

The effect of the age and ethnicity of the different actors on participant recognition rates was not explored because each actor was provided with a subset of emotions/mental states to portray for each modality, preventing a comparison across actors; furthermore, it would be difficult to disentangle acting experience from age effects. Further investigation into cross-cultural, age, and gender effects on participants' emotion recognition would be worthy of further investigation but was also not possible with the present data set, due to differences in

participant characteristics between the three data collection sites. The gender and age distributions were homogeneous between the surveys within each data collection site, but heterogeneous between sites.

Given the overall high recognition scores and the findings comparable to those from other emotion validation studies from the US, Belgium, Switzerland, and the Netherlands with predominantly younger populations (Bänziger et al., 2012; Goeleven et al., 2008; Langner et al., 2010; Tottenham et al., 2009), suggests this innovative validation design seems to be effective when investigating a large number of emotion/mental state categories. Despite some of the limitations discussed above, the EU-Emotion Stimulus Set is a valuable resource for scientific research into emotion perception across both individual and integrated expression modalities.

References

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science, 16,* 403–410. doi:10.1111/j.0956-7976.2005.01548.x

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva Multimodal Expression corpus for experimental research on emotion perception. *Emotion, 12,* 1161–1179.

Bölte, S., & Poustka, F. (2003). The recognition of facial affect in autistic and schizophrenic subjects and their first-degree relatives. *Psychological Medicine, 33,* 907–915.

Comparelli, A., Corigliano, V., De Carolis, A., Mancinelli, I., Trovini, G., Ottavi, G., & Girardi, P. (2013). Emotion recognition impairment is present early and is stable throughout the course of schizophrenia. *Schizophrenia Research, 143,* 65–69. doi:10.1016/j.schres.2012.11.005

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Frank, M. G., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology, 80,* 75–85.

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: A validation study. *Cognition and Emotion, 22,* 1094–1118. doi:10.1080/02699930701626582

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge Mindreading (CAM) Face–Voice Battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 3,* 169–183.

Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., . . . Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, *115*, 137–143. doi:10.1016/S0165-0270(02)00006-7

Kanade, T., Cohn, J. F., & Tian, Y. (2000, March). *Comprehensive database for facial expression analysis*. Paper presented at the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG 2000). Grenoble, France.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage, 18,* 156–168.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24,* 1371–1388. doi:10.1080/02699930903485076

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010, June). *The Extended Cohn–Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression*. Paper presented at the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, CA, USA.

Lundqvist, D., Berggren S., O'Reilly H., Tal, S., Fridenzon, S., Newman, S., . . . Bölte, S. (2014). *The difficulty and importance of emotions in children with a diagnosed Autism spectrum condition according to parents and clinical experts*. Manuscript submitted for publication.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces—AKDEF* (CD ROM, ISBN 91-630-7164-9). Stockholm, Sweden: Karolinska Institutet, Department of Clinical Neuroscience, Psychology section. Downloadable from www.emotionlab.se

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, 102,* 16518–16523. doi:10.1073/pnas.0507650102

Meletti, S., Benuzzi, F., Rubboli, G., Cantalupo, G., Stanzani Maserati, M., Nichelli, P., & Tassinari, C. A. (2003). Impaired facial emotion recognition in early-onset right mesial temporal lobe epilepsy. *Neurology, 60,* 426–431.

O'Toole, A. J., Harms, J., Snow, S. L., Hurst, D. R., Pappas, M. R., & Abdi, H. (2005). A video database of moving faces and people. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27,* 812–816. doi:10.1109/TPAMI.2005.90

Regenbogen, C., Schneider, D. A., Finkelmeyer, A., Kohn, N., Derntl, B., Kellermann, T., & Habel, U. (2012). The differential contribution of facial expressions, prosody, and speech content to empathy. *Cognition and Emotion, 26,* 995–1014. doi:10.1080/02699931.2011.631296

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Cognitive Brain Research, 20,* 81–91.

Sucksmith, E., Allison, C., Baron-Cohen, S., Chakrabarti, B., & Hoekstra, R. A. (2013). Empathy and emotion recognition in people with autism, first-degree relatives, and controls. *Neuropsychologia, 51,* 98–105. doi:10.1016/j.neuropsychologia.2012.11.013

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168,* 242–249. doi:10.1016/j.psychres.2008.05.006

Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research, 1284,* 100–115. doi:10.1016/j.brainres.2009.05.075

Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7,* 487–494.

Van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. J. (2011). Moving faces, looking places: The Amsterdam Dynamic Facial Expressions Set (ADFES). *Emotion, 11,* 907–920. doi:10.1037/a0023853

Weyers, P., Mühlberger, A., Hefele, C., & Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology, 43,* 450–453. doi:10.1111/j.1469-8986.2006.00451.x