Article

PERCEPTION

Perception
2018, Vol. 47(10–11) 1029–1042

DOI: 10.1177/0301006618797226
journals.sagepub.com/home/pec
SAGE

Development of the Youth Emotion Picture Set

Bruno Novello, Anelise Renner, Guilherme Maurer, Soraia Musse and Adriane Arteche
Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Brazil

Abstract
Several facial expression image sets have been developed. Nevertheless, there is a lack of facial expression sets comprising adolescents' images depicting all basic emotions. This study aimed to fill this gap through the development of an image database of youth facial expressions, containing pictures of six basic emotions plus neutral. Posed and spontaneous expressions were collected from 31 youths, 12 to 20 years old; 2,279 frames were obtained, and an initial screening was conducted through the exclusion of similar frames, low intensity images, and ambiguous or blended expressions; 256 frames met criteria and were rechecked by two expert judges. Images were retained if they depicted all the prototypical features of the designated expression. A final selection was conducted to assure an image set that covered all age ranges, both sexes and an even number of images by expression, resulting in 42 frames (21 male, six of each emotion). Expert judges, independent adults, independent teenagers, and a software validation were used to assure database validity. Agreement across raters was high, and no differences were observed for posed and spontaneous images. The data set developed in this study can be a valid tool in studies of facial expressions, in particular, with adolescents' samples.

Keywords
facial expression, emotions, adolescence, affect recognition

Date Received: 15 September 2017; accepted: 6 August 2018

Introduction

Facial expressions have played an important role in human survival because they allow emotional messages, such as fear in the presence of danger, or disgust at perceived risk of

Corresponding author:
Adriane Arteche, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga 6681, Building 11/9 Floor, RS 90619-900, Porto Alegre, Brazil.
Email: adriane.arteche@pucrs.br

food contamination, to be sent at high speed without the need of verbal cues (Fridlun, 1994). Even in less hostile habitats than those in which primitive men lived, facial expressions continue to play an important role in human life, particularly because they present the potential to improve the quality of social interactions (Suzuki, Hoshino, Shigemasu, & Kawamura, 2007). This is remarkably important during adolescence, when the greatest changes in social interactions occur (Brown, 2004) and the ability to identify facial expressions becomes even more fundamental for proper socialization (Thomas, De Bellis, Graham, & Labar, 2007). Nevertheless, the mechanisms underlying facial affect recognition and the developmental trajectory of facial recognition during childhood and adolescence are not yet fully understood (Baird et al., 1999; Batty & Taylor, 2006; Ganel, Valyear, Gottstein, & Goodale, 2005).

Methodological inconsistencies such as various stimuli (e.g., different technical qualities and features) and multiple task designs might account for results' heterogeneity (Langner, Dotsch, Bijlstra, & Wigboldus, 2010). In addition, due to the lack of available high-quality stimuli of youth, many studies with children and teenagers' samples have used adult image databases (McClure, 2000), largely neglecting the so-called own age bias (Anastasi & Rhode, 2005; Wright & Stroud, 2002). Biasing effects suggest that schemas for emotional face recognition are based upon our more recent and significant relationships and that we are, therefore, better at recognizing faces of individuals who mimic our own age and social characteristics (He, Ebner, & Johnson, 2011). These assumptions lead to a key flaw in facial recognition studies that use children and adolescent samples, but adult stimuli (Rhodes & Anastasi, 2012).

Despite its striking relevance, validated image databases comprising images not only of adults but also of children and adolescents are still lacking. The most widely image sets used for facial recognition tasks are the *Pictures of Facial Affect* and the *Japanese and Caucasian Facial Expression of Emotion* (Olszanowski et al., 2015). Common sets recently developed depict adult images, for example, the *Warsaw Set of Emotional Facial Expressions Pictures*, composed of 210 images of 30 models aged 20 to 30 years (Olszanowski et al., 2015); the *Karolinska Emotional Directed Faces Database* (490 images of 70 models, aged 20–30 years; Goeleven, Raed, Leyman, & Verschuere, 2008); and the *Nimstim* (672 frames of 43 professional actors aged 21–30 years; Tottenham et al., 2009). The *Radboud Faces Dataset* differs from other sets of images presenting adult (aged 17–22 years) and children (aged 8–12 years) photographs expressing the six basic emotions, the same photos taken from five different angles, and each picture with three gaze direction's variation (Langner et al., 2010). Few databases comprise pictures over a wider age span and even those do not cover the full adolescence span or do not cover all basic emotions. Those databases are the *NIMH Child Emotional Faces Picture Set* (Egger et al., 2011), which does not include pictures of disgust and surprise expressions and is comprised of 534 photographs of subjects aged 10 to 17 years; the *Dartmouth Database of Children Faces*, which is composed of photographs of Caucasian children aged 5 to 16 years (80 models, both sexes split evenly; Dalrymple, Gomez, & Duchaine, 2013); and the *FACES*, comprised of photographs from models ranging in age from 21 to 80 years (Ebner, Riediger, & Lidenberger, 2010)—therefore not including younger teenagers.

To achieve valid images of emotional expressions, researchers have used multimodal strategies, such as interaction with actors and repetition of phrases (Gross & Levenson, 1995). The use of emotional scenarios to obtain facial expressions was first reported in a study that aimed to investigate anxiety expressions (Perkins, Inchley-Mort, Pickering, Corr, & Burgess, 2012). The researchers followed the *Questionnaire of Treat Response* norms (Blanchard, Hynd, Minke, Minemoto, & Blanchard, 2001) to create brief narratives in

which the listener had to imagine himself or herself in and pose the expression related to the situation. All discrete emotions proposed by Ekman (1992) were contemplated in this instrument. Those scenarios were adapted to Brazilian Portuguese, and the narratives were recorded in MP3 files (Borges, 2013). An additional technique is the use of video stimuli to elicit emotions and therefore obtain expressions. There is a high degree of ecological advantage in the use of this strategy even though there is not a set of videos widely accepted by the scientific community (Gross & Levenson, 1995). The categorization of potential videos to be used with this aim has already been attempted (Leupoldt et al., 2007; Rottenberg, Ray, & Gross, 2007; Scaini et al., 2017) but not yet validated for teenagers' samples in Brazil.

Due to these limitations, the use of posed expressions (PEs) has been the standard procedure to produce emotional frames (LoBue & Thrasher, 2014). In fact, there is a disadvantage in the use of spontaneous expressions in comparison to posed facial expressions given that spontaneous expressions in dynamic contexts are expected to have differences in the speed, trajectory, symmetry, and order of muscle groups' activation (Valstar, Pantic, Ambadar, & Cohn, 2006). Hess and Kleck (1994) demonstrated that the recognition of happiness and disgust statistically differ between spontaneous and PEs, with spontaneous expressions having greater decoding accuracy. However, no significant difference was observed for the other expressions.

Considering the lack of image sets covering all teenagers' age range, this study aimed to develop the *Youth Emotion Picture Set*, an image set comprising images of youth aged 12 to 20 years, depicting six basic emotions plus neutral unambiguously. For the development of this image set, three different strategies were employed: emotive scenarios (ESs), reaction (RE) to visual stimuli, and PEs. The study methodology is in line with gold standards for the development of databases of emotional facial expressions (e.g., Tottenham et al., 2009). Two steps were followed (a) database development (image acquisition, selection of frames, and frames' treatment) and (b) database validation (expert ratings, independent ratings, and software validation). Procedures and results of each step are described later.

Method

Database Development

Image acquisition. Participants. Thirty-one volunteers (14 males) recruited using a snowball method served as models for the images that compose the instrument. Participants age ranged from 12 to 20 years ($M = 17.4$, $SD = 2.7$) and were split into four age groups (12–14 years, $n = 5$; 15–16 years, $n = 6$; 17–18 years, $n = 8$; and 19–20 years, $n = 12$). Regarding racial background, participants were self-declared Caucasian, Black ($n=1$), or mixed ($n = 3$). The exclusion criterion was the presence of facial features that could be distracting, such as scars.

Procedures. Image acquisition took approximately 2 h and 30 min and was conducted at the university laboratory, or in the participants' house according to their preference. Legal guardians of minor age participants and participants aged 18 years or more provided written informed consent. An informed assent term was also required from <18 participants. Researchers responsible for this study signed a term of responsibility for the use of the images. A 1080 HD camera and a tripod were used to record images. A 2-m high black cloth was fixed in the walls where the filming happened, and volunteers wore a 1-m long black cape to hide clothes. Models were required to remove large jewels but were allowed to retain small earrings. Recording of facial expressions was performed in three phases.

Initially, the volunteer was left alone in the room with an MP3 audio playing emotional scenario audios (to avoid possible embarrassment), and he or she was required to imagine himself or herself in 23 situations while making the corresponding expression. Stimuli and instructions used in this part followed Borges' (2013) procedures.

The emotional scenarios were presented in a scrambled order (the same procedure used by Borges, 2013), lasted 16 to 21 s and included situations such as: "You are having a great time with your friends" (happiness); "You are walking alone in the night, when a cyclist approaches, pulls a gun and announces a robbery" (fear); "You are focused in an activity and suddenly an object falls from the shelf and makes a great noise" (surprise); "You are ready to fight" (anger); "You have just discovered that you stepped on dog poop" (disgust); and "You are suffering a lot because you ended a relationship" (sadness). Next, three images of each one of the six basic emotions (happiness, anger, surprise, fear, disgust, and sadness) extracted from the *NimStim Image Set* (Tottenham et al., 2009) displaying different emotion intensities (15%, 50%, and 100%), plus eight images developed in the flash tool ARTNATOMY or ARTNATOMIA (two for fear, one for each one of the other five basic emotion, and a neutral image; Flores, 2005), were used as templates to help the volunteers to pose the expressions. Volunteers were asked to mimic the given expression. To encourage participants, the researcher also mimics the expression immediately after.

Finally, volunteers were exposed to six short movies (1–5 min long, for the emotions happiness, disgust, sadness, fear, surprise, and one base line movie) and to a game (to elicit anger). As proposed by Rottenberg et al. (2007), a baseline video was presented for as long as the previous stimuli to allow participants to get back to their emotional baseline state. Each video was selected by the first author, a member of the research team, and it was previously tested in a pilot study with 10 participants (a boy and a girl of each age range 12–14, 15–16, 17–18, and 19–20 years). In the pilot study, participants were required to watch the video and choose which emotion they were mostly feeling using a six alternative forced choice: fear, happiness, disgust, anger, sad, and surprise. All participants mentioned the intended emotion regardless of their age. Videos were presented in a fixed order: surprise, sadness, fear, disgust, anger, and happiness. Stimuli order was fixed with surprise as the first one that participants were more susceptible to this kind of RE, once that they could not use deductive reasoning to identify what was the target emotion, and because it could mislead to happiness stimuli since it begins as a prank video. In addition, because the disgust stimuli could lead to intrusive thoughts, as identified during the pilot study, it was followed by the anger stimuli since the game leads to a broadly cognitive engagement. The happiness stimuli were selected to be the last one, in order to allow participants to finish the experiment in a positive mood.

A special attention was given to anger and fear as these are the most difficult emotions to elicit in research (Gross & Levenson, 1995), usually requiring personal involvement to be elicited (Rottenberg et al., 2007). For anger, a flash game with elevated difficulty and cheap tricks was used. During the pilot study, it was identified that six trials were enough to prompt this emotion. For sadness, an excerpt of the movie *The Champ* was used since it had already been used in emotional research (Rottenberg et al., 2007). With the exception of anger and sadness, the remaining emotions (surprise, disgust, happiness, and fear) and the baseline stimuli were taken from the website *YouTube* and included videos with the following contents: Happiness: An expert from an auditory TV showing where the participant and the host have communication difficulties; Fear: The first person perspective short movie called *One Last Dive* in which a police diver finds a ghost under water; Disgust: A short movie where an inflamed sebaceous cyst is squeezed; and Surprise: A short video called "Roommate Alien Prank Goes Bad" that starts as a prank video with an alien theme but

ends with a real attempt of abduction. None of the volunteers had previously seen the movies, but three volunteers reported previous contact with the game.

Results. In all, 2,279 frames were extracted using the three applied strategies: ESs ($n = 707$), PEs ($n = 1{,}224$), and REs ($n = 348$). Selected frames comprised 1,175 males (ES: 434/PE: 541/RE: 200) and 1,104 females (ES: 273/PE: 683/RE: 148) images. These frames were screened by the first author and those that were very similar or that depicted clearly ambiguous expressions were removed (first filtering). A total of 727 frames (395 males, ES: 113/PE: 238/RE: 36/, and 332 females, ES: 97/PE: 199/RE: 25) were retained. Figure 1 shows the fluxogram of frames' acquisition by acquisition strategy and emotion.

Frames' selection. *Participants and procedures.* Three expert judges (all psychologists, with expertise in cognitive psychology and trained in emotion facial recognition via the *Facial Action Coding system* developed by Ekman (Ekman & Friesen, 1978; golden rule 80% in accuracy) were asked to exclude: (a) images with illumination issues (e.g., over illumination and side lighting), (b) asymmetric face angles, and (c) blended or ambiguous expressions. Aiming to standardize the criteria and to eliminate ambiguous and blended expressions, judges were required to analyze three facial muscle groups: (a) forehead and eyebrows, (b) eyes, and (c) mouth. To retain a picture, at least two of the muscle groups should demonstrate coherent and similar contractions to those proposed by Ekman and Friesen (2003), and none of the muscle groups could depict an ambiguous or incoherent contraction (one muscle group neutral contraction did not result in picture exclusion).

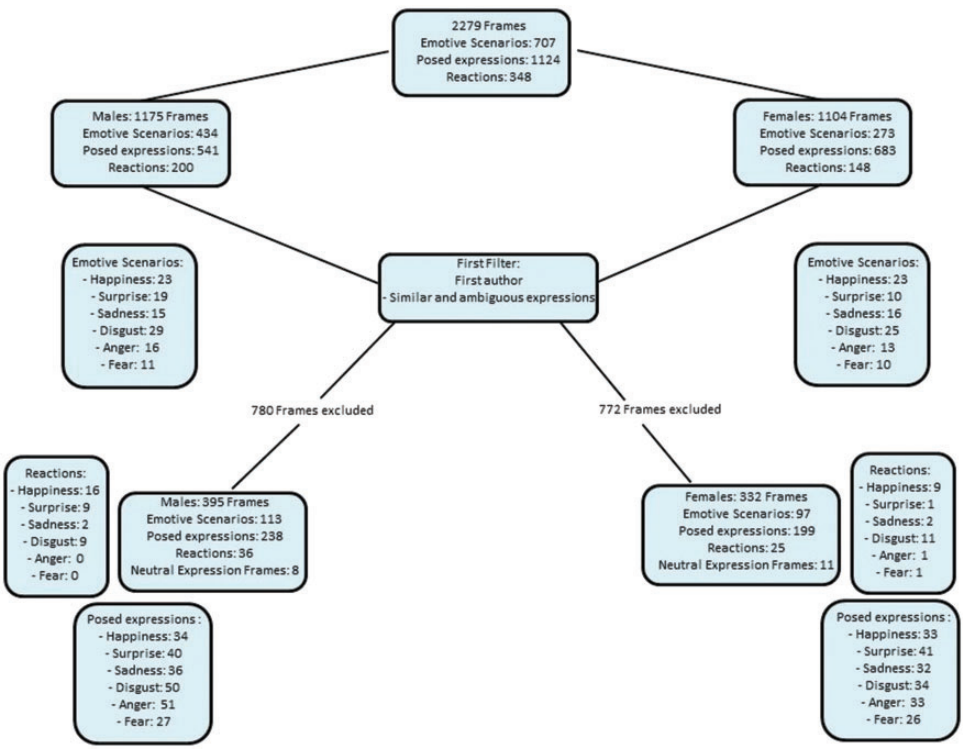

Figure 1. Fluxogram of frames' acquisition.

Results. A total of 256 frames (115 males, ES: 12/PE: 103/RE: 0/; 141 females, ES: 6/PE: 135/RE) were independently retained by the three judges. Then, the same judges selected the three most prototypical frames for each sex, emotion intensity, and racial background to compose the database. A total of 42 images reached >60% agreement and were retained in the database. Pictures retained after this second filtering depicted 21 male images (ES: 3/PE: 15/RE: 0/, 3 neutral) and 21 female images (ES: 3/PE: 15/RE: 0/, 3 neutral), and 19 volunteers had their frames selected (8 males). One adolescent had five frames selected, three had four frames selected, two had three frames selected, six had two frames selected, and seven had just one frame retained. The final frames are Caucasian ($n = 37$), mixed, or Black ($n = 5$) and age of 12 to 14 ($n = 5$), 15 to 16 ($n = 7$), 17 to 18 ($n = 12$), 19 to 20 ($n = 18$) years. Figure 2 displays the fluxogram of judges' frame retention. There was no significant difference in the percentage of frame retention between pictures acquires via emotional scenarios and those that we posed ($\chi^2(1) = 2.91$, $p = .08$). Considering the total number of frames extracted from each acquisition method, 3% of those acquired via emotional scenarios and 7.7% of posed frames were retained in the final version. None of the pictures acquired via RE elicitation was retained in the final version.

Image preparation. Procedures. After the selection of the frames, piercings and bracers were removed with Photoshop CS6 (Adobe Systems, 2012). To minimize color effects, all frames were edited as black and white, and the size of the face, eye position, and background were standardized. For this, rulers were placed in a template in Photoshop, splitting the face into three parts (forehead, eyes or nose or mouth, and chin). Subjects' irises and nose were lined up with the rulers, so that way, the eyes' positions in all frames were the same height. Since a black cloth was used in the photographs' background, no changes were necessary on background editing. All images were resized to 369×475 pixels and had the light controlled.

Results. As shown in Figure 3, all images were successfully treated.

Database Validation

Expert judges. Participants. Four expert judges (see frames' selection section for a full description of judges' background) rated the selected frames according to the emotion displayed (happiness, sadness, anger, disgust, fear, surprise, and neutrality).

Procedures. Frames were randomly presented to the judges using an online platform. Expert judges did not have set time to rate the picture and were asked to select one out of the seven forced-choice emotion options (fear, happiness, disgust, anger, sad, surprise, neutral). Data collection was online, and participants were required to use a notebook or a desktop with a screen 15″ to perform the task.

Results. Expert judges' ratings reached 100% of agreement for all the pictures displaying anger, disgust, surprise, fear, and neutrality. Only two pictures, one displaying a happy face, and the other displaying a sad face were mistakenly rated by one of the judges, thus achieving 75% of agreement for those emotions. The happy face was considered a neutral emotion, and the sad face was rated as disgust. The overall Kappa coefficient revealed high agreement (Kappa: 0.972). There was no difference on average agreement of posed versus spontaneous frames ($p > .10$).

Independent judges—Adults. Participants. In all, 101 independent adults judges aged between 18 and 77 years (age: 18–25 years [42.6%]; 26–39 years [26.6%], and 40 years or more [20.8%]), 68.3% were females, most of them having completed high school (41.6%) or

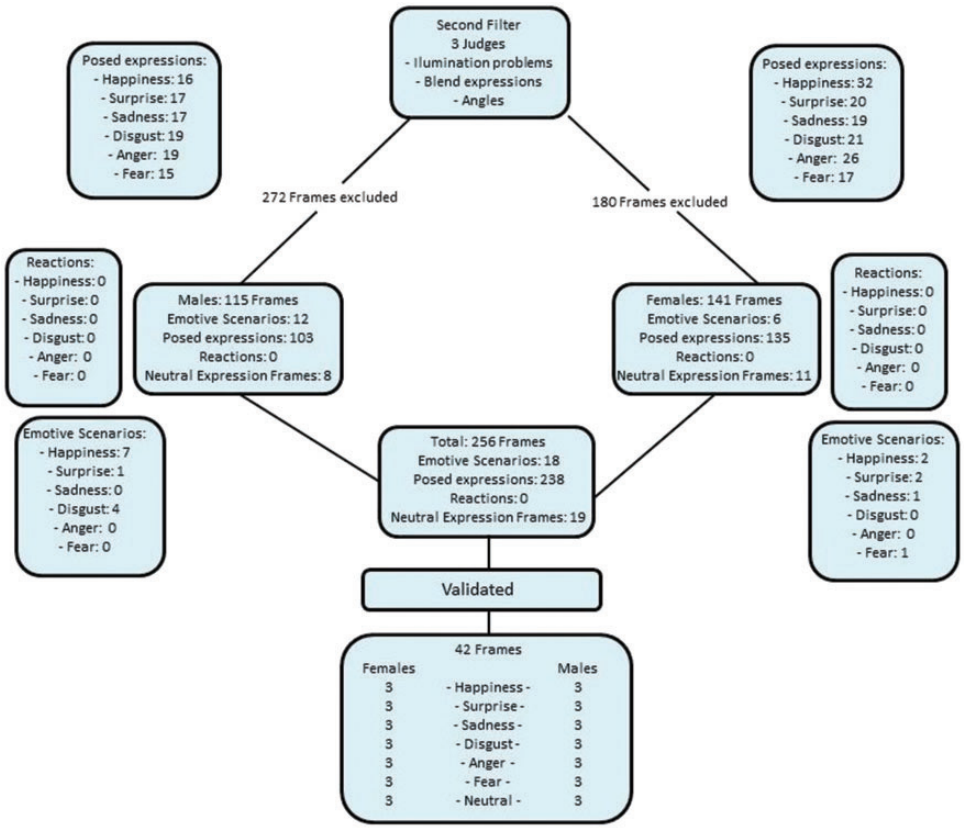

Figure 2. Fluxogram of judges' frame retention.

postgraduate (34.7%), only 21.8% were psychologists, not trained to recognize facial expressions were asked to rate the 42 pictures' emotion.

Procedures. Participants were recruited via a snowball method. Data collection was performed online, and procedures followed those of expert judges. Judges' agreement was calculated per frame. Average agreement and standard deviations of judges' agreement for each emotion were also computed.

Results. Faces 3, 4, 23, 24, 28, 32, 33, and 34 showed 100% agreement across judges. Faces 1, 2, 5, 6, 13, 19, 20, 21, 22, 25, 26, 27, 29, 30, 36, and 39 had above 95% agreement. The remaining frames had the following agreement percentages: Frame 7 (82.2%), Frame 8 (79.2%), Frame 9 (64.4%), Frame 10 (72.3%), Frame 11(88.1%), Frame 12 (69.4%), Frame 14 (94.1%), Frame 15 (78.2%), Frame 16 (91.1%), Frame 17 (89.1%), Frame 18 (49.5%), Frame 31(91.1%), Frame 35(55.4%), Frame 37 (92.1%), Frame 38 (36.6%), Frame 40 (81.2%), Frame 41 (80.2%), and Frame 42 (85.1%). Average agreement by emotion ranged between 75.9% (fear) and 98% (happiness). Standard deviation of independent adults' agreement ranged from $SD = 2.35$ (happiness) to $SD = 19.43$ (sadness). There was no difference in the average agreement of posed versus spontaneous frames ($p > .10$).

The sex of the judges had no effect on pictures' ratings except for Frame 42, $\chi^2(2) = 10.748$, $p = .005$, that 93.7% of men identified as a sad face, but 81.1% of women did so.

Figure 3. Final image data set.
P = posed; ES = emotional scenario; H = happy; F = fear; N = neutral; D = disgust; A = anger; Su = surprise; Sa = sad.

The two men who did not attribute sadness rated it as a disgust image, and all women who attributed another emotion for the face rated it as fear. Judges' age had a significant effect on ratings of Frames 9, $\chi^2(4) = 10.570$, $p = .032$: 55.8% of the judges aged 18 to 25 years, 83.8% of the 26 to 39 year olds, and 47.6% of those older than 40 years identified as fear; Frame 12, $\chi^2(4) = 11.320$, $p = .023$: 55.8% of the judges aged 18 to 25 years, 78.4% of the 26 to 39 year olds, and 81.0% of those older than 40 years identified as fear. Judges who did not attribute fear for these two images rated it as surprise; and Frame 19, $\chi^2(2) = 7.095$, $p = .029$: 88.4% of the judges aged 18 to 25 years, 100% of the 26 to 39 year olds, and 100% of those older than 40 years identified it as disgust, and the others rated it as anger.

Significant effects of education were observed in Frames 9, $\chi^2(6) = 59.831$, $p = .000$: 47.6% of the judges with complete high school, 73.9% of the judges with complete higher education, and 80.0% with postgraduation identified it as fear. Two judges attributed disgust, and the others identified it as surprise; Frame 40, $\chi^2(9) = 19.604$, $p = .021$: 81.0% of the judges with complete high school, 79.3% of the judges with complete higher education, and 88.6% with postgraduation identified it as sadness. The others identified it as fear, disgust, or anger. When comparing psychologists with other professionals, significant differences were found in Frame 9, $\chi^2(2) = 15.578$, $p < .001$, which 100% of the psychologists and 54.4% of the other professionals identified as fear (those who mistakenly rated the face attributed either surprise or disgust).

Independent judges—Adolescents. Participants. A total of 54 adolescents aged 12 to 17 years ($M = 14.95$, $SD = 1.60$, 59.3% female) and without prior training in face recognition participated as judges.

Procedures. Participants were recruited in state schools. Informed consent was obtained from all caregivers, and all teenagers signed assent terms. Data collection was carried out individually, in person, and at school. Stimuli presentation procedures followed those of expert and independent-adult judges, except that adolescents had a fixed 1000 ms to rate the picture.

Results. Frame 2 showed 100% agreement across judges. The other frames had the following agreement's rates: 7(51.9%), 8(44.4%), and 1, 3, 4, 5, 6, 13, 14, 16, 20, 22, and 9(61.1%), 10 (46.3%), 11(70.4%), 12(40.6%), 15(79.6%), 17(92.6%), 18(88.9%), 19(88.9%), 21(83.3%), 23(96.3%), 24(94.4%), 25(85.2%), 26(87.0%), 27(85.5%), 28(92.6%), 29(85.2%), 30 (81.5%), 31(88.9%), 32(94.4%), 33(90.7%), 34(96.3%), 35(59.3%), 36(90.7%), 37(60.8%), 38(63.0%), 39(94.4%), 40(70.4%), 41(66.7%), and 42(70.4%). Average agreement by emotion ranged from 52.4% (fear) to 87% (anger). Standard deviation of adolescents' agreement ranged between $SD = 3.75$ (anger) and $SD = 15.90$ (sadness). There were no significant effects of judges' sex or age on images' ratings (all $p > .05$), and there was no difference in the average agreement between posed and spontaneous frames ($p > .10$).

Software validation. Procedures. Images were uploaded to the Microsoft Emotion API (https://westus.dev.cognitive.microsoft.com/docs/services/5639d931ca73072154c1ce89/operations/563b31ea778daf121cc3a5fa) which identified a face rectangle in each image and returned a percentage of certainty for each one of the eight target emotions (anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise). Even though we had not designed, the study to include contempt faces we opted to retain this emotion in the software validation process as we considered evidence of validity if none of the faces fell into the contempt category. We also opted to retain the disgust emotion, albeit its algorithm is still experimental in the API.

Results. As shown in Table 1, all but one set of images were identified in accordance with the ground truth meaning that the correct emotion was the one identified with the highest score. Disgust was the only emotion which was not correctly identified by the software; however, this was expected since such emotion is still experimental in the API. Standard deviation of software's emotion correct attribution ranged between $SD = 0.04$ (happiness) and $SD = 10.07$ (disgust).

Finally, average ratings of frames were aggregated by emotion and compared across the three applied raters (independent adults, independent adolescents, and software) using repeated measure analysis. Bonferroni correction was applied for pairwise comparisons. Happiness and surprise showed no significant differences across raters ($p > .06$). A significant effect was found for fear, $F(2, 10)=14.15$, $p=.001$, $\eta^2=.74$, with teenagers being significantly less accurate than adults ($p=.012$). A rater effect was also found for neutral faces, $F(2, 10)=3.92$, $p=.055$, $\eta^2=.44$, with adolescents being less accurate than the API software ($p=.034$). Similarly, in sadness, $F(2, 10)=5.45$, $p=.025$, $\eta^2=.52$, adolescents showed a tendency to be less accurate than the API ($p=.062$). For anger, $F(2, 10)=11.99$, $p=.002$, $\eta^2= .71$, teenagers were significantly less accurate than adults ($p=.010$) and the API ($p=.026$). For disgust, $F(2, 10)=52.80$, $p<.001$, $\eta^2=.44$, as expected due to the software limitation, the API was less accurate than teenagers ($p<.001$) and adults ($p<.001$). Figure 4(a) shows the average correct attribution for independent adults, independent adolescents, and the software and (b) shows the difference between the software correct attribution and both independent judges and adolescents.

Table 1. API Software Emotion Identification for Each Frame.

	Anger (%)	Contempt (%)	Disgust (%)	Fear (%)	Happiness (%)	Neutral (%)	Sadness (%)	Surprise (%)
1	0%	0%	0%	0%	100%	0%	0%	0%
2	0%	0%	0%	0%	100%	0%	0%	0%
3	0	0	0	0	100	0	0	0
4	0	0	0	0	100	0	0	0
5	0	0	0	0	100	0	0	0
6	0	0	0	0	100	0	0	0
7	6	0	0	86	0	0	0	6
8	31	0	1	64	0	0	0	2
9	12	0	0	58	0	0	0	27
10	0	0	0	63	0	0	0	36
11	11	0	4	77	0	0	4	1
12	2	0	0	70	0	0	0	25
13	0	0	0	0	0	99	0	0
14	0	0	0	0	0	99	0	0
15	0	0	0	0	0	96	3	0
16	0	0	0	0	0	98	1	0
17	0	0	0	0	0	99	0	0
18	0	0	0	0	0	98	0	0
19	83	0	16	0	0	0	0	0
20	61	0	29	1	0	0	6	0
21	1	0	32	0	64	0	1	0
22	43	2	42	0	0	9	1	0
23	50	1	41	1	0	2	2	1
24	1	0	18	2	0	0	77	0
25	89	0	10	0	0	0	0	0
26	99	0	0	0	0	0	0	0
27	99	0	0	0	0	0	0	0
28	99	0	0	0	0	0	0	0
29	89	0	1	0	0	0	7	0
30	99	0	0	0	0	0	0	0
31	0	0	0	3	2	0	0	93
32	0	0	0	0	0	0	0	99
33	0	0	0	0	0	0	0	99
34	0	0	0	0	0	0	0	99
35	0	0	0	1	14	0	0	84
36	0	0	0	0	0	0	0	100
37	0	0	2	1	0	0	96	0
38	1	0	3	2	0	13	76	1
39	0	0	0	0	0	6	93	0
40	0	0	2	0	0	0	97	0
41	0	0	0	0	0	0	99	0
42	0	0	1	0	0	0	97	0

Discussion

This study aimed to develop a database of adolescents' emotional expressions. For the development of this image set, 31 youths aged 12 to 20 years agreed to provide pictures of their facial expressions depicting the six basic emotions plus neutral. To acquire the

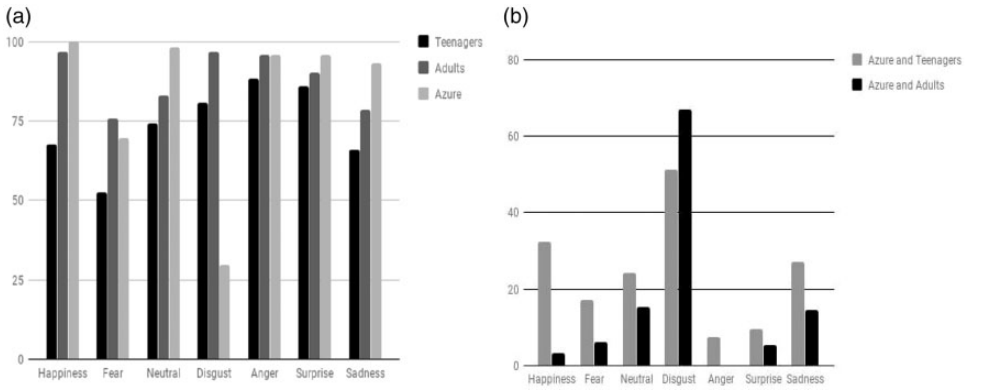

Figure 4. Average correct attribution for lay adults, lay adolescents, and software and difference between the software correct attribution and lay judges.

images, we employed different strategies: ESs, RE to visual stimuli, and PEs. We used independent adolescents and adults as well as a software analysis to validate the images. Our final database comprises 42 (six of each emotion and six neutral faces) validated images of boys and girls of different racial backgrounds. ESs and PEs were effective strategies to achieve good quality images, nevertheless none of the images acquired via RE to visual stimuli was retained in the final data set.

The posed procedure was the most effective method to acquire the frames, which can be justified by the fact that with this method, there is immediate feedback of the expressions produced since it is possible for the researcher to correct blended expressions. The emotional scenarios' strategy created fewer images. Nevertheless, the proportion of images from emotional scenarios and PEs retained in the final version was not significantly different meaning that those elicited via emotional scenarios were potentially very representative of the target emotion. It should be considered that in this modality, subjects were instructed to act naturally watching the videos; therefore, some frames could not be captured because of the gaze direction, and changes in the head angle.

Racial diversity was a difficult criterion to meet. Only four participants from the database were not White, and the final database retained 14% of multiracial and Black volunteers' pictures. As participants aged 17 to 18 years were better at posing expressions, there was a higher prevalence of this age group in the database, which should be considered in studies with younger samples. Also, the lowest amount of excluded female frames from the PEs strategy was a result of them being more expressive, requiring fewer attempts. The ESs strategy appears to be influenced by demographic variables, such as sex, and the volunteers' personal experience. For example, some of the participants began to laugh during the sad scenario (e.g., "You are suffering a lot because you ended a relationship"), and the female participants were more prone to make sad expressions in the anger scenario (e.g., "You found someone who you hate making gossips about you"). It should be noted that in the study conducted by Borges (2013), which also used those scenarios, participants were all adult actors, different from our study in which the volunteers were adolescents with no previous acting experience.

The final 42 images retained were subject to three different raters: independent adults, independent teenagers, and a software validation. Overall, agreement was high and in the expected direction, with adults having a better accuracy and very similar to the software and teenagers being the least accurate raters. This is in line with the developmental trajectory of

face recognition that shows that the ability to recognize emotional faces is associated with hormonal levels and will not be complete until late adolescence (Guapo et al., 2009). In addition, differences in methodological procedures might account for part of this variation given that adolescents were given 1000 ms to view the image, whereas adults had no set time. Previous studies have shown that even 200 ms is a sufficient exposure to allow conscious attribution of a set emotion (Schyns, Petro, & Smith, 2009), nevertheless, a limited exposure time may have influenced the slightly lower accuracy ratings of teenagers.

The image set developed in this study can be used in research aiming the study of facial expressions recognition in adolescents enabling the control of a possible age of stimuli effect, as well as in experimental tasks related to face recognition. As the expressions selected to compose the image set had high intensity as criteria, we advise that they would be used in tasks with different time exposures in order to prevent a ceiling effect. Furthermore, this database can be used in research with clinical application purposes, through the development of a training program in facial expressions recognition for adolescents with social anxiety disorder, for example.

Despite its strengths and potential applications, this study has some limitations. First, videos used to elicit spontaneous emotions were the same for younger and older teenagers. Another limitation is the fact that REs were the last strategy used, so there may be an effect of the volunteers being tired, suggesting that the results could differ if participants had been exposed to the three strategies in a random order. In addition, it is worth noting that judges in the study used forced-choice options to identify faces' emotions, and this might have inflated accuracy rates.

As a suggestion for future studies, it is possible to use the ESs to determine if specific characteristics, such as clinical disorders, are related to a greater expressiveness in certain types of ESs. Also, the number of posing individuals could be increased allowing phenotypical diversity. It should also be considered that the high agreement across judges was observed for all emotions possibly due to the careful filtering process in previous steps, as well as the standardized training of the judges. In addition, the selection of frames was made with the objective of selecting prototypal expressions and if low intensity expressions were included, the agreement rate could be less favorable. Finally, a multimodal approach could favor the selection of more ambiguous, nonprototypical pictures to be used in a wider variety of settings.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Bruno Novello has received financial support for his/her master degree course (from which this research was derived) from *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES).

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional or national research committee, and with the 1964 Helsinki declaration and its later amendments, or comparable ethical standards. Informed consent was obtained from all individual participants included in the study. As some participants were minors, their legal guardians

received an informed consent term where the research objectives, procedures, the image consent for usage in scientific events, and for the purpose of research were described. To confirm whether the minors were fully aware of the objectives, procedures, and future image use, an informed assent term was also required to be signed by them. In addition, the researchers signed a statement of responsibility for the use of the images.

References

Adobe Systems. (2012). *Adobe photoshop CS6 user guide*. San Jose, CA: Adobe Systems Inc.

Anastasi, J., & Rhodes, M. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*, *12*, 1043–1047. doi:10.3758/BF03206441

Baird, A., Gruber, S., Fein, D., Maas, L., Steingard, R., Renshaw, P., . . . Yurgelun-Todd, D. (1999). Functional magnetic resonance imaging of facial affect recognition in children and adolescent. *Journal of Child and Adolescent Psychiatry*, *38*, 195–199. doi:10.1097/00004583-199902000-00019

Batty, M., & Taylor, M. (2006). The development of emotional face processing during childhood. *Developmental Science*, *9*, 207–220. doi:10.1111/j.1467-7687.2006.00480.x

Blanchard, D., Hynd, A., Minke, K., Minemoto, T., & Blanchard, R. (2001). Human defensive behaviors to threat scenarios show parallels to fear- and anxiety-related defense patterns of non-human mammals. *Neuroscience and Biobehavioral Reviews*, *25*, 761–770. doi:10.1016/S0149-7634(01)00056-2

Borges, V. F. (2013). *Expressão facial de ansiedade: Análise no julgamento de observadores* [Anxiety expression.: Raters´ agreement analysis] (master's Dissertation). University of Londrina, Londrina, Brazil.

Brown, B. B. (2004). Adolescents' relationships with peers. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 363–394). Hoboken, NJ: John Wiley & Sons, Inc.

Dalrymple, K., Gomez, J., & Duchaine, B. (2013). The Dartmouth database of children's faces: Acquisition and validation of a new face stimulus set. *PLoS One*, *8*, e79131. doi:10.1371/journal.pone.0079131

Ebner, N., Riediger, M., & Lidenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*, 351–362. doi:10.3758/BRM.42.1.351

Egger, H., Pine, D., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K., & Angold, A. (2011). The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): A new set of children's facial emotion stimuli. *International Journal of Methods in Psychiatric Research*, *20*, 145–156. doi:10.1002/mpr.343

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*, 169–200. doi:10.1080/02699939208411068

Ekman, P., & Friesen, W. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologist Press.

Ekman, P., & Friesen, W. (2003). *Unmasking the face*. Los Altos, CA: Malor Books.

Flores. (2005). *Artnatomy*. Retrieved from www.artnatomia.net

Fridlun, A. (1994). *Human facial expression: An evolutionary view*. Cambridge, MA: Academic Press.

Ganel, T., Valyear, K., Goshen-Gottstein, Y., & Goodale, M. (2005). The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia*, *43*, 1645–1654. doi:10.1016/j.neuropsychologia.2005.01.012

Goeleven, E., Raed, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition & Emotion*, *22*, 1094–1118. doi:10.1080/02699930701626582

Gross, J., & Levenson, R. (1995). Emotion elicitation using films. *Cognition & Emotion*, *9*, 87–108. doi:10.1080/02699939508408966

Guapo, V., Graeff, F., Zani, A., Labate, C., Reis, C., & Del-bem, C. (2009). Effects of sex hormonal levels and phases of the menstrual cycle in the processing of emotional faces. *Psychoneuroendocrinology*, *34*, 1087–1094. doi:10.1016/j.psyneuen.2009.02.007

He, Y., Ebner, N. C., & Johnson, M. K. (2011). What predicts the own-age bias in face recognition memory? *Social Cognition*, *29*, 97–109.

Hess, U., & Kleck, R. (1994). The cues decoders use in attempting to differentiate emotion-elicited and posed facial expressions. *European Journal of Social Psychology*, *24*, 367–381. doi:10.1002/ejsp.2420240306

Langner, O., Dotsch, R., Bijlstra, G., & Wigboldus, D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*, 1377–1388. doi:10.1080/02699930903485076

Leupoldt, A., Rohde, J., Beregova, A., Thordsen-Sörensen, I., Zur Nieden, J., & Dahme, B. (2007). Films for eliciting emotional states in children. *Behavior Research Methods*, *39*, 606–609.

LoBue, V., & Thrasher, C. (2014). The Child Affective Facial Expression (CAFE) set: Validity and reliability from untrained adults. *Frontiers in Psychology*, *5*, 1–8.

McClure, E. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infant, children and adolescents. *Psychological Bulletin*, *126*, 424–453. doi:10.1037/0033-2909.126.3.424

Olszanowski, M., Pochwatko, G., Kuklinski, K., Scibor-Rylski, M., Lewinski, P., & Ohme, R. K. (2015). Warsaw set of emotional facial expression pictures: A validation study of facial display photographs. *Frontiers in Psychology*, *5*, 1516. doi:10.3389/fpsyg.2014.01516

Perkins, A., Inchley-Mort, S., Pickering, A., Corr, P., & Burgess, A. (2012). A facial expression for anxiety. *Journal of Personality and Social Psychology*, *102*, 910–924. doi:10.1037/a0026825

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, *138*, 146.

Rottenberg, J., Ray, R., & Gross, J. (2007). Emotion elicitation using films. In. J. A. Coan & J. J. Allen (Eds.), *The handbook of emotion elicitation and assessment* (pp. 9–28). London, England: Oxford University Press.

Scaini, S., Rancoita, P. M V., Martoni, R. M., Omero, M., Ogliari, A., & Brombin, C. (2017). Integrating dimensional and discrete theories of emotions: A new set of anger- and fear-eliciting stimuli for children. *The Journal of Genetics Psychology*, *178*, 253–261. doi:10.108/00221325.2017.1351416

Schyns, P. G., Petro, L. S., & Smith, M. L. (2009). Transmission of facial expressions of emotion co-evolved with their efficient decoding in the brain: Behavioral and brain evidence. *PLoS One*, *4*, e5625. doi:10.1371/journal.pone.0005625

Suzuki, A., Hoshino, T., Shigemasu, K., & Kawamura, M. (2007). Decline or improvement? Age-related differences in facial expression recognition. *Biological Psychology*, *74*, 75–84. doi:10.1016/j.biopsycho.2006.07.003

Thomas, L., De Bellis, M., Graham, R., & LaBar, K. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science*, *10*, 547–558. doi:10.1111/j.1467-7687.2007.00614.x

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

Valstar, M., Pantic, M., Ambadar, Z., & Cohn, J. (2006). Spontaneous vs. posed facial behavior: Automatic analysis of brow actions. In *Proceedings of ACM International Conference on Multimodal Interfaces* (pp. 162–170). Retrieved from https://doi.org/10.1145/1180995.1181031

Wright, D., & Stroud, J. (2002). Age differences in lineup identification accuracy: People are better with their own age. *Law and Human Behavior*, *26*, 641–654. doi:10.1023/A:1020981501383