

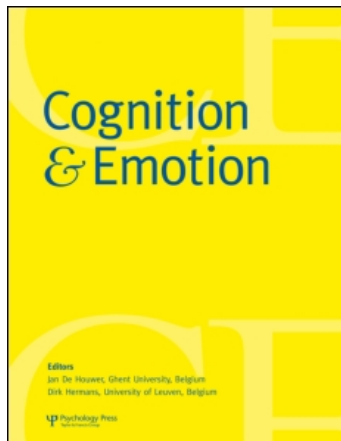
This article was downloaded by: [Radboud University Nijmegen]

On: 24 November 2010

Access details: Access Details: [subscription number 907172236]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Cognition & Emotion

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713682755>

Presentation and validation of the Radboud Faces Database

Oliver Langner^a; Ron Dotsch^a; Gijsbert Bijlstra^a; Daniel H. J. Wigboldus^a; Skyler T. Hawk^b; Ad van Knippenberg^a

^a Radboud University Nijmegen, Nijmegen, The Netherlands ^b University of Amsterdam, Amsterdam, The Netherlands

Online publication date: 22 November 2010

To cite this Article Langner, Oliver , Dotsch, Ron , Bijlstra, Gijsbert , Wigboldus, Daniel H. J. , Hawk, Skyler T. and van Knippenberg, Ad(2010) 'Presentation and validation of the Radboud Faces Database', Cognition & Emotion, 24: 8, 1377 – 1388

To link to this Article: DOI: 10.1080/02699930903485076

URL: <http://dx.doi.org/10.1080/02699930903485076>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Presentation and validation of the Radboud Faces Database

Oliver Langner, Ron Dotsch, Gijsbert Bijlstra, and Daniel H. J. Wigboldus

Radboud University Nijmegen, Nijmegen, The Netherlands

Skyler T. Hawk

University of Amsterdam, Amsterdam, The Netherlands

Ad van Knippenberg

Radboud University Nijmegen, Nijmegen, The Netherlands

Many research fields concerned with the processing of information contained in human faces would benefit from face stimulus sets in which specific facial characteristics are systematically varied while other important picture characteristics are kept constant. Specifically, a face database in which displayed expressions, gaze direction, and head orientation are parametrically varied in a complete factorial design would be highly useful in many research domains. Furthermore, these stimuli should be standardised in several important, technical aspects. The present article presents the freely available Radboud Faces Database offering such a stimulus set, containing both Caucasian adult and children images. This face database is described both procedurally and in terms of content, and a validation study concerning its most important characteristics is presented. In the validation study, all frontal images were rated with respect to the shown facial expression, intensity of expression, clarity of expression, genuineness of expression, attractiveness, and valence. The results show very high recognition of the intended facial expressions.

Keywords: Face database; Validation; Emotion; Gaze direction.

Face processing may well be one of the most complex tasks that human beings accomplish. Faces carry a wealth of social information, including information about identity, emotional and motivational status, lip speech, and focus of attention as indicated by eye gaze, all of which are important for successful communication. Not surprisingly, face

images are commonly used as stimulus materials in a wide range of research fields such as, for instance (the development of) face processing, emotion research, information processing in phobias and autism, social referencing, interpersonal attraction, persuasive communication, person memory, impression formation, human and computer face

Correspondence should be addressed to: Oliver Langner, Behavioural Science Institute, PO Box 9104, NL-6500 HE Nijmegen, The Netherlands. E-mail: o.langner@psych.ru.nl

The Radboud Faces Database was financed by the Behavioural Science Institute of the Radboud University Nijmegen. Skyler T. Hawk is currently a postdoctoral researcher at the Research Centre Adolescent Development, Utrecht University.

We thank Job van der Schalk for his help as a FACS specialist.

recognition, and robotics (e.g., Calder & Young, 2005; Grossmann & Johnson, 2007; Rule, Ambady, Adams & Macrae, 2008; Schwaninger, Wallraven, Cunningham, & Chiller-Glaus, 2006).

Different fields naturally put varying demands on the specific characteristics of the stimulus materials they use. For example, studies on facial expression processing mainly use straight-gaze images with standardised expressions (e.g., Vuilleumier, Armony, Driver, & Dolan, 2003), whereas research on facial attention cues uses face images with varying gaze directions and head orientations (e.g., Loomis, Kelly, Pusch, Bailenson, & Beall, 2008). Research on the development of face-processing abilities requires similar stimuli based on children's faces (e.g., Guyer et al., 2007). Furthermore, recent research indicates that interactions between different facial-stimulus characteristics yield additional social information. For example, it has been found that gaze direction affects the perception of emotions (Adams & Kleck, 2005), and that attractive faces enhance evaluations of associated products when their gaze is directed towards the participant but not when it is averted (Strick, Holland, & van Knippenberg, 2008). Unfortunately, researchers often rely on different stimulus sets, ad hoc assembled stimuli from the internet, or even faces more or less successfully edited by image software, because the currently available facial databases contain only manipulations of specific stimulus characteristics and typically not of combinations of characteristics. Moreover, the facial stimuli used in research tend to vary substantially in terms of technical features and overall technical quality. As a result, diverging findings from different studies may be attributable to variations in the facial or technical characteristics of the stimulus sets used.

In our view, a face database with parametric variations on a number of important facial characteristics constitutes an important extension of the currently available stimulus materials. Crucial characteristics, with a wide applicability in different fields of research, are facial expression, gaze direction, and head orientation, as well as a reasonable number of male and female models of both

adults and children. Furthermore, the stimuli included in this database should be controlled for potentially interfering technical factors like positions of facial landmarks, lighting conditions, and image background. See Table 1 for an overview of often used databases and their features.

In the present article, we present the freely available Radboud Faces Database (RaFD), a face database containing Caucasian face images that vary along all previously mentioned characteristics and provide adequate control of the indicated technical factors. All models in the dataset show eight facial expressions with three gaze directions, photographed simultaneously from five different camera angles. The photos were taken in a highly controlled environment. All displayed facial expressions were based on prototypes from the *Facial Action Coding System* (FACS; Ekman, Friesen, & Hager, 2002a). As an important addition, the dataset contains images of both adult and child models, all with the same stimulus characteristics and the same level of technical control.

In the following report, we present validation data for all frontal camera images. For the validation, 276 participants rated all images of nine randomly chosen models from the database. For each image, participants rated the depicted facial expression, and the intensity, clarity, valence and genuineness of the expression. Further, they rated the attractiveness of the neutral frontal gaze images for all nine models. This enables an assessment of the quality of the dataset. Furthermore, researchers can use these ratings to select images from the dataset with the specific properties they require. Mean ratings for all individual images and rating dimensions are available as online support material.

DEVELOPMENT OF DATABASE

Description of the image set

The database contains portrait images of 49 models in two subsets: 39 Caucasian Dutch adults (19 female), and 10 Caucasian Dutch children (6 female). All models showed eight facial expressions with three gaze directions (see Figure 1a and b).

Table 1. *Examples of existing face databases and their features*

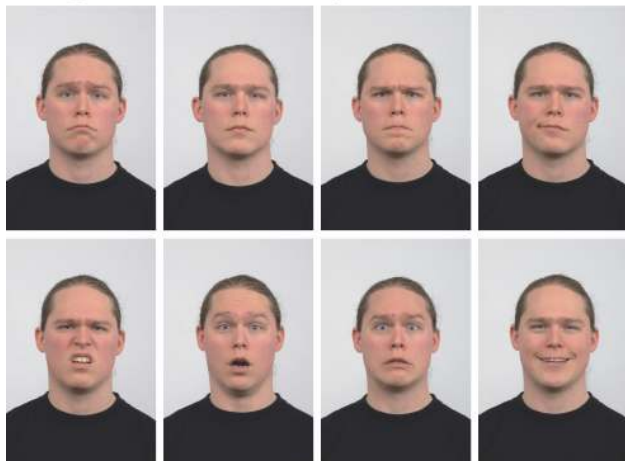
<i>Name of set</i>	<i>Authors</i>	<i>Models</i>	<i>Features</i>
JACFEE/JACNeuF	Matsumoto & Ekman (1988)	56	Female and male models Japanese and Caucasian models Seven emotions and neutral expression All images frontal gaze and 90° camera
KDEF	Lundqvist, Flykt, & Öhman (1998)	70	Female and male models, all Caucasian Six emotions and neutral expression Frontal gaze Five camera angles
MSFDE	Beaupré, Cheung, & Hess (2005)	12	Female and male models French Canadian, Chinese, and sub-Saharan African models Six emotions and neutral expression All images frontal gaze and 90° camera
IAPS	Lang, Bradley, & Cuthbert (1999)	—	Female and male models Different ethnicities Neutral, sad, angry images in various contexts
ADFES	van der Schalk, Hawk, & Fischer (2009)	20	Female and male models Caucasian and Turkish/Moroccan models Nine emotions, dynamic expressions Two gaze directions
Facial Expression Subset	Hawk, van Kleef, Fischer, & van der Schalk (2009)	8	Female and male models Nine emotions and neutral expression Dynamic expressions All frontal gaze and 90° camera
NimStim	Tottenham et al. (1998)	45	Female and male models Caucasian, Latin American, African American, and Asian American models Seven emotions and neutral expression Frontal gaze

Expressions were neutral, anger, sadness, fear, disgust, surprise, happiness, and contempt, the expressions most consistently recognised across cultures (Ekman, 1992). Each expression was shown with eyes directed straight ahead, averted to the left, and averted to the right. Photos were taken against a uniform white background from five different camera angles simultaneously, with viewpoints from left to right in steps of 45° (see Figure 1c). This amounts to 120 images per model. Models wore black t-shirts, had no hair on the face and wore no glasses, makeup or jewellery. The targeted emotional expressions were based on prototypes defined in the *Investigator's Guide for the Facial Action Coding System* (Ekman, Friesen, & Hager, 2002b). The action

units we targeted, using a variation of the Directed Facial Action Task (e.g., Ekman, 2007), are shown in Figure 2.

The photo shoot took place in early 2008 at Radboud University Nijmegen. Beforehand, models practiced all emotional expressions at home for at least one hour, following a detailed training manual. Throughout the photo shoot, two certified FACS specialists coached all models. During the session, each model first practised all expressions with a FACS specialist for 25 minutes. The actual photo shoot took about 45 minutes for each model, during which one of the FACS specialists monitored the expressions on a TV screen while giving detailed instructions for each expression.

a) Eight Emotional Expressions



b) Three Gaze Directions



c) Five Camera Angles



Figure 1. Examples from the Radboud Faces Database. (a) Examples for the eight emotional expressions. From top left: sad, neutral, angry, contemptuous, disgusted, surprised, fearful, happy. (b) Examples for the three gaze directions. (c) Examples for the five camera angles. Here cameras in the order: 180°, 135°, 90°, 45°, 0°. [To view this figure in colour, please visit the online version of the paper.]

Apparatus

We used five Nikon cameras (models D200, D2X, and D300), with resolutions between 10 and 12 Mpx. Three 500 W flashes (Bowens Int., Essex) were used for illumination. All cameras and flashes were connected to a wireless remote control (Pulsar, Bowens Int., Essex), allowing us to take photos on all cameras simultaneously (see Figure 3 for the technical setup).

Image processing

All photos were initially stored in raw format. Photos were converted to tiff-image format and corrected for white-balance by using the free software packages UFRaw and The Gimp. Next, all images were spatially aligned according to facial landmarks using Matlab (The Mathworks Inc., Natick, MA). We

used a simplex optimisation algorithm for this purpose, fitting two translational, and one rotational alignment parameters. First, we aligned each image of a model to the neutral straight gaze image of the same model and camera angle. Then, all neutral straight gaze images of the different models were aligned towards each other. Finally, both within- and between-models alignment parameters were applied to the images, and all aligned images were cropped and resized to a size of 1024 × 681 pixels.

VALIDATION OF DATABASE

Method

Participants and apparatus. A total of 276 students (238 female) from the Radboud University

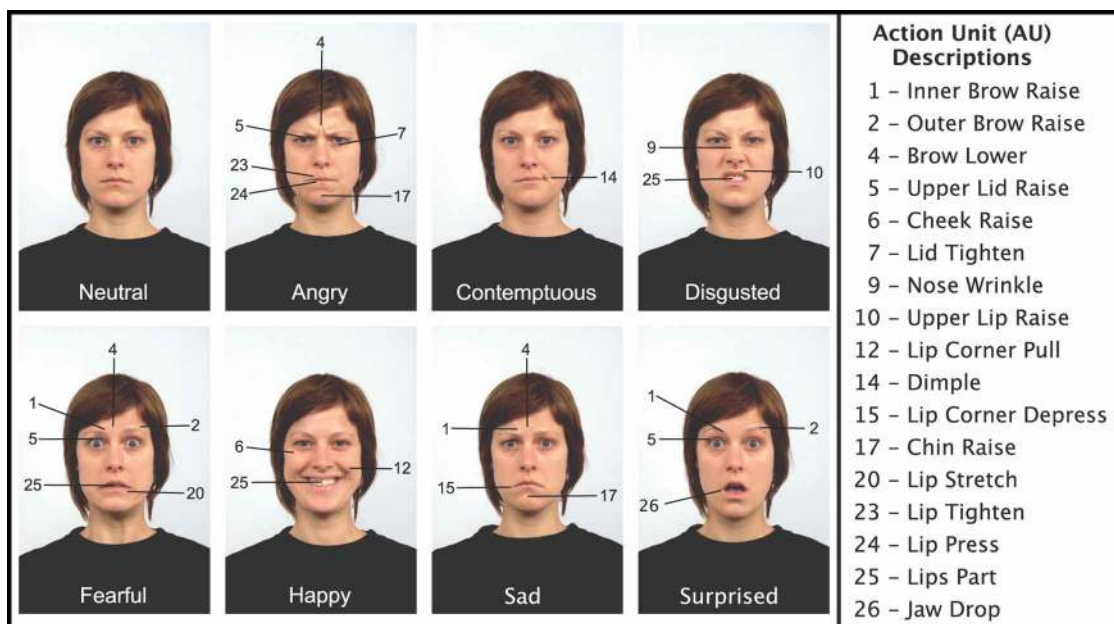


Figure 2. Targeted action units (AU) for all emotional expressions. [To view this figure in colour, please visit the online version of the paper.]

Nijmegen participated in the validation study, with a mean age of 21.2 years (SD 4.0). All had normal or corrected-to-normal vision and received €10 or course credits for participation. The validation study was programmed using PsychoPy (Peirce, 2007), a python library for conducting psychological experiments.

Procedure. Only the frontal view images (90° -camera) were validated. Participants were presented with pictures from only one of the database subsets, either adults or children. For each model, gaze direction, and facial expression, originally two images were present in the validation stimulus set. From these two only the one with the highest inter-rater agreement concerning the intended expression was retained for inclusion in the final database.

The validation started with an attractiveness rating. Participants scored the neutral, straight-gaze images of all subset models on a 5-point scale, ranging from unattractive to attractive. Image

order was randomised. This task familiarised the participants with all models in the subset.

Next, participants rated the images of 9 subset models on several dimensions. For each model, participants viewed images with all three gaze directions combined with eight emotional expressions, summing up to 216 images for each participant. This way, participants saw equal numbers of emotions from each model. Which models were presented was chosen randomly across participants, with the constraint that every image was rated by at least 20 participants.

In each trial, a randomly chosen image from the 9 subset models presented to the participant was shown in the centre of the screen. For each image, participants successively judged: (a) the depicted expression; (b) the intensity of the expression; (c) the clarity of the expression; (d) the genuineness of the expression; and (e) the valence of the image, in this order. Before the task, participants got instructions for each rating dimension. For each judgement, the current

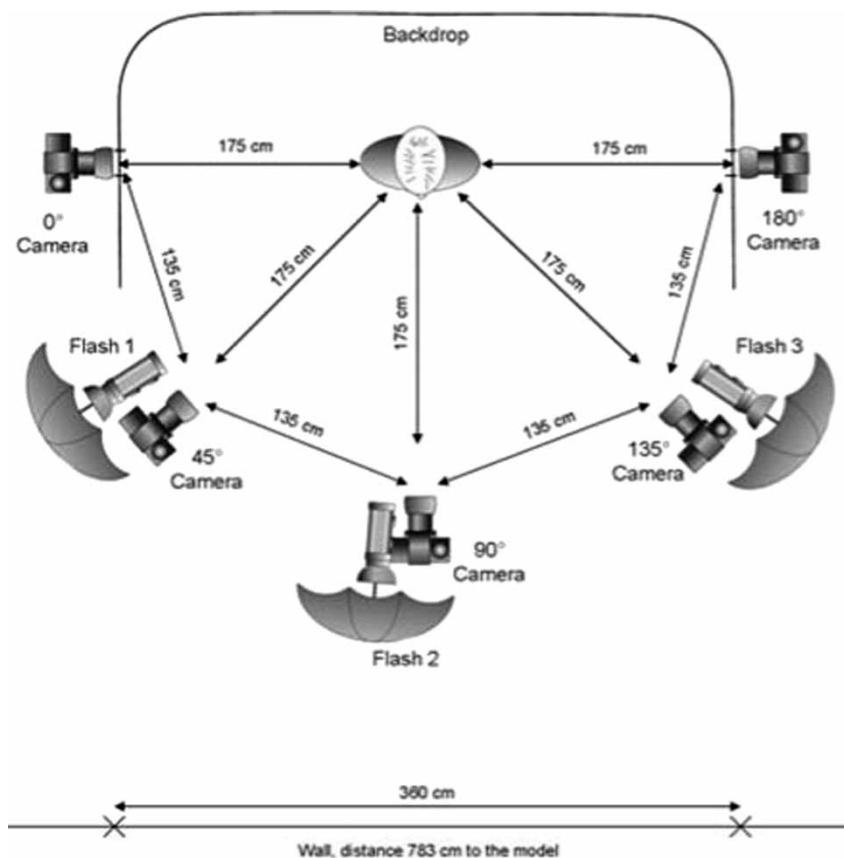


Figure 3. Technical setup of the photo shoot.

rating dimension was indicated above the image (e.g., “Clarity”) and the corresponding rating scale was displayed below it. The expression rating was forced-choice with nine response categories, i.e., the eight expressions used in the dataset and “other” (Frank & Stennett, 2001). We asked participants to pick the emotion label that best fitted the shown facial expression. The ordering of the expression labels from which participants could choose was counterbalanced across participants, but kept constant within participants. All other dimensions were rated on 5-point scales. We instructed participants to rate the *emotional expression* of the shown face with regard to the intensity (“weak” to “strong”), the clarity (“unclear” to “clear”), and the genuineness (“faked” to “genuine”) of the expression. Finally, we asked

participants to judge the overall valence of the image (“negative” to “positive”).

Results

Each image was rated at least 20 times on each rating dimension. For the measures of expression, intensity, clarity, genuineness, and valence, separate analyses of variance (ANOVAs) were computed with the factors Subset (children, adults), Gender (male, female), Expression (neutral, anger, sadness, fear, disgust, surprise, happiness, contempt), and Gaze Direction (frontal, left, right). Due to the high number of images, most statistical tests were significant. Therefore only effects with an $\eta^2_p > .10$ are reported, corresponding to $p < .01$. Means and SDs of all measures are shown for each subset and expression

in the appendix. Indices for inter-rater reliability are reported in Table 2 for all appropriate rating dimensions.¹ Further, on-line supporting materials containing average validation data of individual images are available. Please go to www.rafd.nl.

Attractiveness. Mean attractiveness ratings (*SDs*) were 2.36 (0.53) and 2.10 (0.58) for the female and male adult models, respectively; and 2.42 (0.40) and 2.44 (0.63) for the female and male child models, respectively.

Expression. For each image, we calculated how many participants chose the targeted emotion. Overall the agreement between chosen and targeted emotions was 82% (median 88%, *SD* 19%). Average choice rates per targeted emotion are depicted in Figure 4. An ANOVA of the arcsine-transformed agreement rates (Winer, 1971) revealed a significant effect of expression, $F(7, 1080) = 168.2, p < .01, \eta^2_p = .52$. Post hoc tests showed that agreement was significantly higher for happiness (mean 98%, *SD* = 3%), and significantly lower for contempt (mean 50%, *SD* = 15%), compared to all other expressions (means between 80 and 90%).

Table 2. Intraclass correlations for all rating dimensions (but the emotion rating)

Dimension	ICC(1, 1)	ICC(1, k)
<i>Dutch adult</i>		
Attractiveness	.31	.99
Intensity	.20	.83
Clarity	.19	.83
Genuineness	.13	.75
Valence	.44	.94
<i>Dutch children</i>		
Attractiveness	.24	.94
Intensity	.26	.88
Clarity	.22	.85
Genuineness	.09	.67
Valence	.48	.95

¹ Here we report the intraclass correlations ICC(1, 1) and ICC(1, k) as reliability indices (Shrout & Fleiss, 1979). Due to the fact that our participants did not rate the whole image set but only parts, we could not calculate the usually higher indices ICC(2, 1) and ICC(2, k) that partial between-rater variance out. Therefore, our reliability indices are probably lower than the real reliability.

A close look at Figure 4 reveals that, for some expressions, off-diagonal responses were not equally distributed across chosen expressions: Faces with intended surprise were sometimes confused with fear (7%), and, vice versa: intended fear was sometimes confused with surprise (8%); intended disgust was sometimes mistaken for either anger (7%) or contempt (8%); and for intended contempt participants responded frequently either other (24%) or neutral (12%).

As the above-reported agreement rates do not take response bias into account, we additionally determined and analysed unbiased hit rates (Wagner, 1993). Unbiased hit rates can vary between 0 and 1, where a hit rate of 1 indicates not only that a stimulus category (e.g., happy faces) was always identified correctly (e.g., that happy faces were always categorised as happy), but additionally that the corresponding response (e.g., response “happy”) was always used correctly (e.g., that the response “happy” was only given for happy faces). Lower unbiased hit rates result if either stimuli from a category are not classified correctly (e.g., happy faces categorised as angry) or if the corresponding response was also used for stimuli from other categories (e.g., that the response “happy” was also given to angry or surprised faces). Unbiased hit rates were computed as follows: Per participant and for each gaze direction, we first created a choice matrix with targeted and chosen expressions as rows and columns, respectively. Next, the number of ratings in each cell was squared and divided by the product of the marginal values of the corresponding row and column, yielding the unbiased hit rate. As usual for proportions, those were arcsine transformed prior to analysis (Winer, 1971), although for readability untransformed proportions are reported. A repeated-measures ANOVA, with Subset as a between-subjects factor and Model Gender, Gaze, and Expression as within-subjects factors yielded significant effects of Model Gender, $F(1, 274) = 52.2, p < .01$,

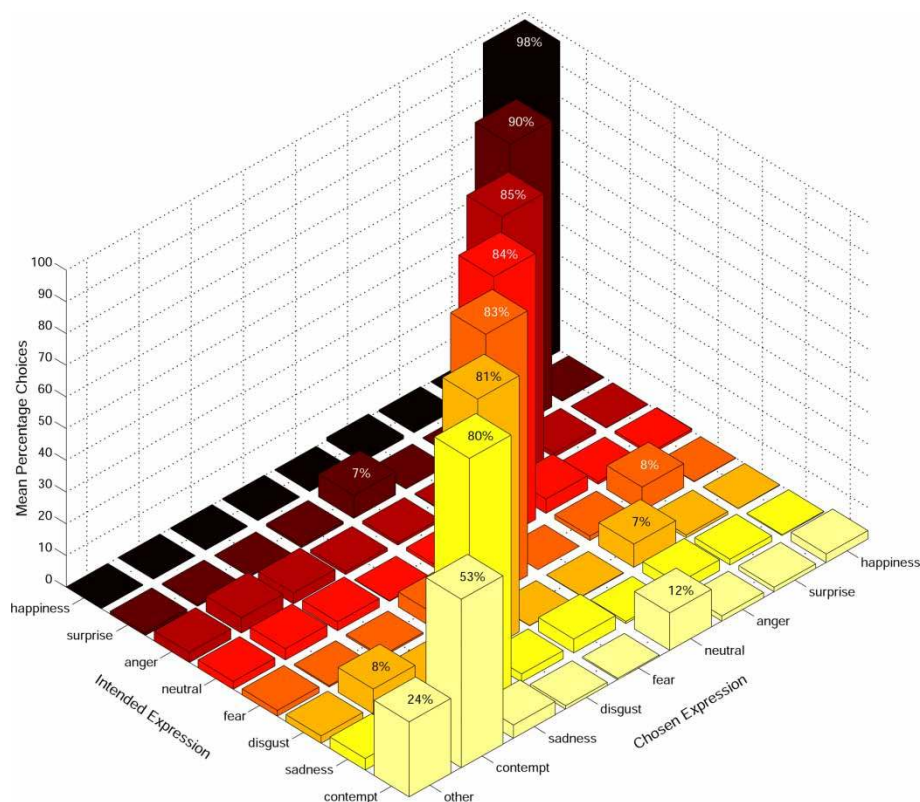


Figure 4. Percentage of chosen emotions per intended emotional expression. [To view this figure in colour, please visit the online version of the paper.]

$\eta^2_p = .16$; and of Expression, $F(7, 1918) = 238.9$, $p < .01$, $\eta^2_p = .47$. Unbiased hit-rates were higher for female than male models, with means (*SDs*) of 73% (27%) and 69% (27%), respectively. Similar to the agreement analysis, post hoc tests revealed significantly higher hit rates for happiness (mean 79%, $SD = 34\%$) and lower hit rates for contempt (mean 29%, $SD = 31\%$) than all other expressions (means between 56 and 65%).

Other validation measures. For each image, we calculated the mean judgments for clarity, intensity, genuineness, and valence. For all measures, ANOVAs showed similar significant expression effects, $F(6, 945) \geq 101.8$, $p \leq .01$, $\eta^2_p \geq .39$. The pattern of results is described for the different validation measures more closely below.

The means of the ratings on the four judgmental dimensions listed above are displayed in the appendix. As can be seen in the appendix, the patterns of means of the judgements of the different expressions are remarkably similar across the Adult and Child datasets. Note that both for adults and children, the ratings on intensity and clarity were highly correlated ($r = .75$), suggesting that the more intense an expression was, the more clear it appeared to be. For instance, happiness, fear, surprise and disgust are both relatively intense and clear, while specifically contempt seems to score low on both measures. Whether this reflects something about the inherent qualities of these expressions, or about the way in which our models expressed them, cannot be assessed on the basis of the present data.

The dimension of genuineness appears to be relatively independent from both intensity and clarity, with correlations of .10 and .24, respectively. Both for the Adult and the Child set, neutral (respectively, $M = 3.9$, $SD = 0.3$; and $M = 3.7$, $SD = 0.3$) and happy faces (respectively, $M = 3.8$, $SD = 0.5$; and $M = 3.5$, $SD = 0.4$) were scored as fairly genuine, while all other expressions scored on average around the nominal midpoint of the scale (i.e., around 3). Finally, the valence of virtually all expressions was rated as expected, with happiness as the only clearly positive expression ($M = 4.2$, $SD = 0.3$ in both sets). Neutral turned out to be truly neutral (Adult set: $M = 3.0$, $SD = 0.3$; Child set: $M = 3.0$, $SD = 0.2$), while surprise was rated fairly close to neutral ($M = 2.8$, $SD = 0.2$ in both sets). All other emotions but contempt were clearly negative ($1.9 \leq M \leq 2.2$), while contempt was slightly less negative than the other negative emotions (Adult set: $M = 2.5$, $SD = 0.2$; Child set: $M = 2.4$, $SD = 0.3$).

DISCUSSION

In this paper we have introduced the Radboud Faces Database, a new database of Caucasian faces with parametric variations of the stimulus characteristics facial expression, gaze direction, and head orientation. RaFD contains images of both adult and child models with the same characteristics and control over technical aspects. We have further reported validation data for all frontal view images.

The overall 82% agreement rate found between intended and chosen expressions was high, lying about 11% beyond that reported in a recent validation study of the Karolinska Directed Emotional Faces database (KDEF; Goeleven, De Raedt, Leyman, & Verschuere, 2008; Lundqvist, Flykt, & Öhman, 1998). The higher median of 88% indicates a skewed distribution of agreement values, with more than 76% of images having an agreement $\geq 80\%$. This underscores the effectiveness of the coaching method employed in eliciting reliable and prototypical facial expressions.

Notably, expression agreement was substantially lower for contempt than for all other expressions. Contempt was also the only expression for which participants frequently chose the response option "other". Although errors during the coaching process for contempt are a possible explanation for this, studies have shown that contempt is a less universally recognised expression across cultures (Ekman & Friesen, 1986; Rosenberg & Ekman, 1995), and that it is subject to dialect-like variations in muscle activations (Elfenbein, Beaupré, Lévesque, & Hess, 2007). Further, Matsumoto and Ekman (2004) found that lower agreement for the expression of contempt reflects problems with the expression label instead of problems with the expression itself. Taken together, this suggests that low agreement on contempt may be a general feature of the emotion, and not of the presented database.

For surprise, fear, and disgust we found systematic patterns of deviating choices. For all three expressions, the most frequently chosen alternative expressions contained morphological overlaps with the intended emotion. These patterns corresponded to those found by Goeleven et al. (2008).

RaFD is a new tool for research using face stimuli. It provides a parametric set of face images varied along important facial characteristics, namely expression, gaze direction, and head orientation. Clearly, there are other facial characteristics important for face processing, like symmetry of faces, masculinity/femininity, distinctiveness, or babyfacedness. Although these have not been systematically varied in the current database, future studies should gather data regarding these characteristics to broaden the applicability of RaFD even further. It should be noted that there is always a trade-off between experimental control and ecological validity of stimuli, and that researchers need to be aware that probably not all factorial combinations of facial characteristics appear equally often in natural environments.

RaFD further contains images from both adults and children, thus being one of the first databases also providing high-quality stimuli for use in developmental research. Importantly, the RaFD is freely available for use in scientific

research. More information about the image set and the application procedure can be found at www.rafd.nl.

Manuscript received 3 June 2009
Revised manuscript received 26 October 2009
Manuscript accepted 3 November 2009
First published online 6 April 2010

REFERENCES

- Adams, R. B., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, 5(1), 3–11.
- Beaupré, M. G., Cheung, N., & Hess, U. (2000). *The Montreal set of facial displays of emotion* [Slides]. (Available from Ursula Hess, Department of Psychology, University of Quebec at Montreal, Montreal, Quebec, Canada.)
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, 6(8), 641–651.
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99(3), 550–553.
- Ekman, P. (2007). The directed facial action task. In *Handbook of emotion elicitation and assessment* (pp. 47–53). Oxford, UK: Oxford University Press.
- Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion*, 10(2), 159–168.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002a). *Facial Action Coding System: The manual*. Salt Lake City, UT: Research Nexus.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002b). *Facial Action Coding System: Investigators guide*. Salt Lake City, UT: Research Nexus.
- Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion*, 7(1), 131–146.
- 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(1), 75–85.
- Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition and Emotion*, 22(6), 1094–1118.
- Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *The European Journal of Neuroscience*, 25(4), 909–919.
- Guyer, A. E., McClure, E. B., Adler, A. D., Brotman, M. A., Rich, B. A., Kimes, A. S., et al. (2007). Specificity of facial expression labeling deficits in childhood psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 48(9), 863–871.
- Hawk, S. T., van Kleef, G. A., Fischer, A. H., & van der Schalk, J. (2009). "Worth a thousand words": Absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion*, 9(3), 293–305.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). *International Affective Picture System (IAPS): Technical Manual and Affective Ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
- Loomis, J. M., Kelly, J. W., Pusch, M., Bailenson, J. N., & Beall, A. C. (2008). Psychophysics of perceiving eye-gaze and head direction with peripheral vision: Implications for the dynamics of eye-gaze behavior. *Perception*, 37(9), 1443–1457.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces—KDEF*. Department of Clinical Neuroscience, Psychology section, Karolinska Institutet.
- Matsumoto, D., & Ekman, P. (1988). *Japanese and Caucasian Facial Expressions of Emotion (JACFEE)* [Slides]. San Francisco, CA: Intercultural and Emotion Research Laboratory, Department of Psychology, San Francisco State University.
- Matsumoto, D., & Ekman, P. (2004). The relationship among expressions, labels, and descriptions of contempt. *Journal of Personality and Social Psychology*, 87(4), 529–540.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.
- Rosenberg, E., & Ekman, P. (1995). Conceptual and methodological issues in the judgment of facial expressions of emotion. *Motivation and Emotion*, 19(2), 111–138.
- Rule, N. O., Ambady, N., Adams, R. B., Jr., & Macrae, C. N. (2008). Accuracy and awareness in the perception and categorization of male sexual orientation. *Journal of Personality and Social Psychology*, 95(5), 1019–1028.
- Schwaninger, A., Wallraven, C., Cunningham, D. W., & Chiller-Glaus, S. D. (2006). Processing of facial

- identity and expression: A psychophysical, physiological, and computational perspective. *Progress in Brain Research*, 156, 321–343.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Strick, M., Holland, R. W., & van Knippenberg, A. (2008). Seductive eyes: Attractiveness and direct gaze increase desire for associated objects. *Cognition*, 106(3), 1487–1496.
- Tottenham, N. (1998). *MacBrain Face Stimulus Set*. Chicago, IL: John D. and Catherine T. MacArthur Foundation Research Network on Early Experience and Brain Development.
- Van der Schalk, J., Hawk, S. T., & Fischer, A. H. (2009). *Validation of the Amsterdam Dynamic Facial Expression Set (ADFES)*. Poster for the International Society for Research on Emotions (ISRE), Leuven, Belgium.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6(6), 624–631.
- Wagner, H. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, 17(1), 3–28.
- Winer, B. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.

APPENDIX

Averages (SDs) of agreement, unbiased hit-rates, and ratings of intensity, clarity, genuineness, and valence per expression and subset separately

<i>Measure</i>	<i>Emotion</i>							
	<i>Anger</i>	<i>Contempt</i>	<i>Disgust</i>	<i>Fear</i>	<i>Happiness</i>	<i>Neutral</i>	<i>Sadness</i>	<i>Surprise</i>
<i>Adults</i>								
Agreement	81 (19)	48 (12)	79 (10)	88 (7)	98 (3)	84 (13)	85 (16)	90 (9)
Unb. Hit-Rate	64 (26)		66 (32)	73 (24)	89 (21)	67 (27)	66 (27)	75 (25)
Intensity	3.5 (0.4)	2.9 (0.3)	3.9 (0.3)	4.1 (0.3)	4.1 (0.4)	3.4 (0.3)	3.4 (0.4)	3.9 (0.3)
Clarity	3.7 (0.4)	2.9 (0.3)	3.9 (0.3)	4.0 (0.3)	4.4 (0.2)	3.6 (0.3)	3.7 (0.4)	4.0 (0.3)
Genuineness	2.9 (0.3)	3.2 (0.2)	3.1 (0.3)	3.0 (0.3)	3.8 (0.5)	3.9 (0.3)	2.9 (0.3)	3.0 (0.3)
Valence	2.0 (0.2)	2.5 (0.2)	2.0 (0.2)	2.1 (0.2)	4.2 (0.3)	3.0 (0.3)	2.1 (0.2)	2.8 (0.2)
<i>Children</i>								
Agreement	89 (14)	59 (19)	83 (18)	79 (13)	97 (6)	84 (16)	75 (25)	91 (8)
Unb. Hit-Rate	61 (23)		64 (26)	55 (25)	91 (14)	69 (25)	63 (25)	71 (22)
Intensity	3.8 (0.5)	3.1 (0.3)	4.1 (0.3)	4.2 (0.4)	3.9 (0.5)	3.2 (0.3)	3.5 (0.4)	4.2 (0.2)
Clarity	3.9 (0.5)	3.1 (0.4)	3.9 (0.4)	4.0 (0.4)	4.2 (0.4)	3.4 (0.4)	3.6 (0.4)	4.1 (0.2)
Genuineness	3.0 (0.3)	3.1 (0.3)	3.2 (0.3)	3.2 (0.3)	3.5 (0.4)	3.7 (0.3)	3.0 (0.4)	3.2 (0.3)
Valence	1.9 (0.2)	2.4 (0.3)	1.9 (0.2)	2.2 (0.2)	4.2 (0.3)	3.0 (0.2)	2.0 (0.2)	2.8 (0.2)