FaceReader
Methodology Note

By Dr. Leanne Loijens and Dr. Olga Krips
Behavioral research consultants at Noldus Information Technology

WHAT IS FACEREADER?

FaceReader is a program for facial analysis. It can detect facial expressions. FaceReader has been trained to classify expressions in one of the following categories: happy, sad, angry, surprised, scared, disgusted, and neutral. These emotional categories have been described by Ekman [1] as the basic or universal emotions. In addition to these basic emotions, contempt can be classified as expression, just like the other emotions [2]. Obviously, facial expressions vary in intensity and are often a mixture of emotions. In addition, there is quite a lot of inter-personal variation.

FaceReader has been trained to classify the expressions mentioned above. It is not possible to add expressions to the software yourself. Please contact Noldus Information Technology if you are interested in the classification of other expressions.

In addition to facial expressions, FaceReader offers a number of extra classifications. It can, for example, detect the gaze direction and whether eyes and mouth are closed or not. With these data you can give an approximation of the test participant's attention.

FaceReader can detect facial expressions: happy, sad, angry, surprised, scared, disgusted, and neutral.
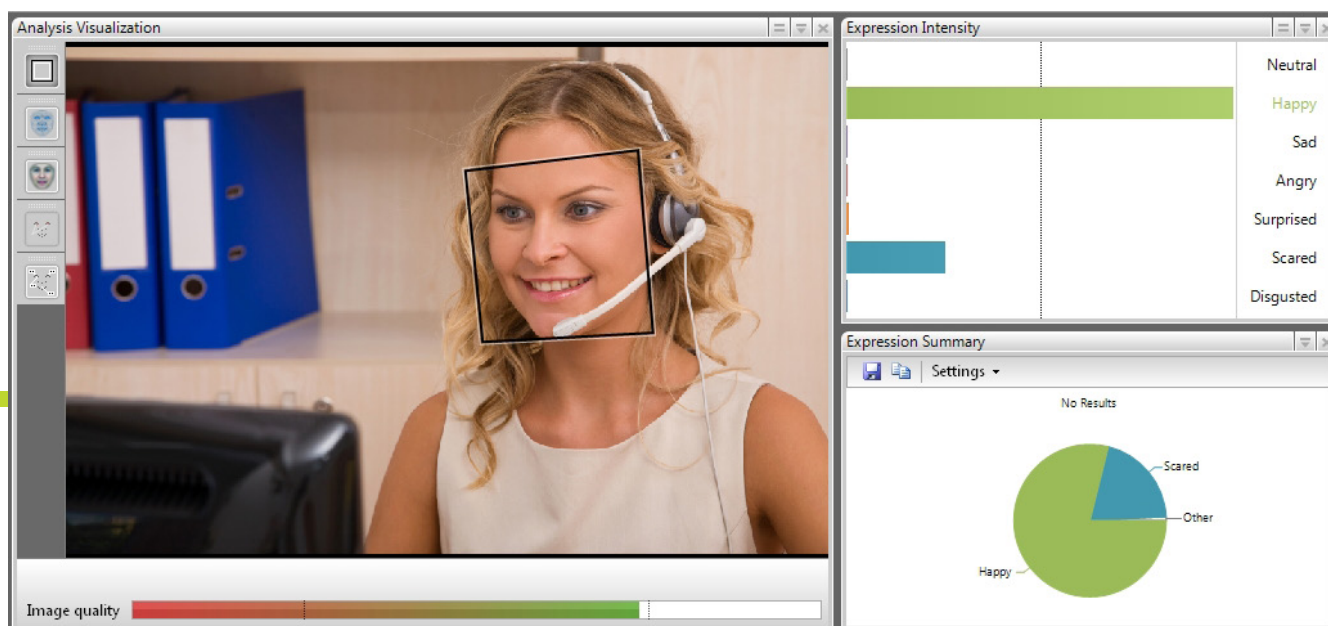
Analyzing facial expressions with FaceReader.

You find a full overview of the classifications in the Technical Specifications of FaceReader that you can obtain from your Noldus IT sales representative. FaceReader can classify facial expressions either live using a webcam, or offline, in video files or images. Depending on the computer you use, Face-Reader can analyze up to 20 frames/second in a live analysis. FaceReader can also record video at 15 frames/second. A prerecorded video can be analyzed frame-by-frame.

HOW DOES FACEREADER WORK?

The Active Appearance Method (AAM) describes over 500 key points in the face and the facial texture of the face, entangled by these points.

FaceReader works in three steps [3,4,5]:

1. The first step in facial expression recognition is detecting the face. Face-Reader uses the popular Viola-Jones algorithm [6] to detect the presence of a face.

2. The next step is an accurate 3D modeling of the face using an algorithmic approach based on the Active Appearance Method (AAM) described by Cootes and Taylor [7]. The model is trained with a database of annotated images. It describes over 500 key points in the face and the facial texture of the face entangled by these points. The key points include

   A. The points that enclose the face (the part of the face that FaceReader analyzes)

   B. Points in the face that are easily recognizable (lips, eyebrows, nose and eyes)

   The texture is important because it gives extra information about the state of the face. The key points only describe the global position and the shape of the face, but do not give any information about, for example, the presence of wrinkles and the shape of the eye brows. These are important cues for classifying the facial expressions.

3. The actual classification of the facial expressions is done by training an artificial neural network [8]. As training material over 10.000 images were used that were manually annotated by trained experts.

With the Deep Face classification method, FaceReader directly classifies the face from image pixels, using an artificial neural network to recognize patterns (so no face finding or modeling is done) [9]. This has the advantage that FaceReader can analyze the face if part of it is hidden. This method is based on Deep Learning, and is done side by side with the Active Appearance

How FaceReader works. In the first step (left) the face is detected. A box is drawn around the face at the location where the face was found. The next step is an accurately modeling of the face (right). The model describes over 500 key points in the face and the facial texture of the face entangled by these points (middle).
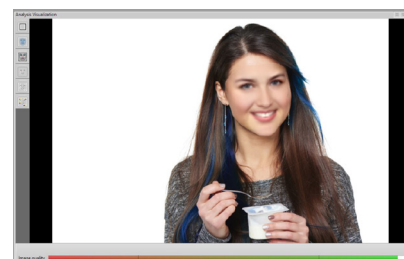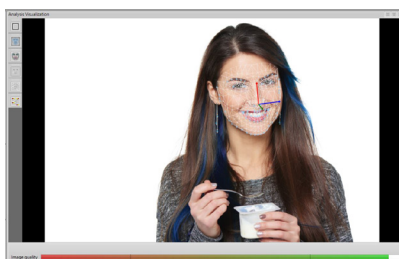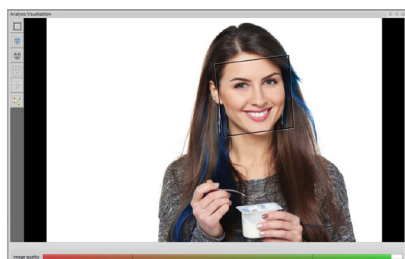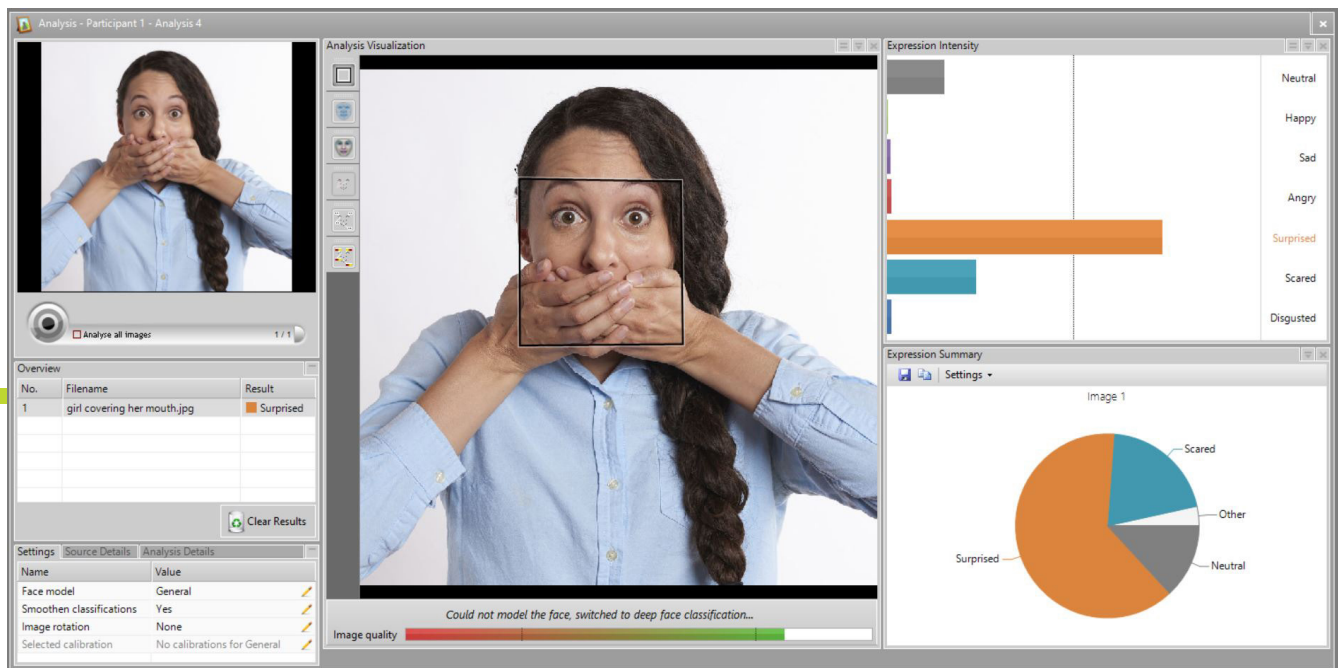
Model and enhances the accuracy of facial expression analysis. In addition to this, Deep Face classification is used stand-alone if modeling with the Active Appearance Model fails, but FaceReader is still able to determine the position of the eyes. In this case, the following analyses can be carried out:

- Facial expression classification
- Valence calculation
- Arousal calculation
- Action Unit classification
- Subject characteristics analysis

There are multiple face models available in FaceReader. In addition to the general model which works well under most circumstances for most people, there are models for East Asian people, elderly, and children. Before you start analyzing facial expressions, you must select the face model which best fits the faces you are going to analyze.

CALIBRATION

For some people, FaceReader can have a bias towards certain expressions. You can calibrate FaceReader to correct for these person-specific biases. Calibration is a fully automatic mechanism. There are two calibration methods, participant calibration and continuous calibration. Participant calibration is the preferred method. However, if you have the project analysis module, do not use either calibration method, but calculate the expressions relative to those during a neutral stimulus instead (see the section *The Project Analysis Module*).

With the Deep Face classification method FaceReader can analyze the face if part of it is hidden.

Calibration is a fully automatic mechanism. There are two calibration methods, participant calibration and continuous calibration.

For Participant calibration, you use images or camera or video frames in which the participant looks neutral. The calibration procedure uses the image, or frame with the lowest model error and uses the expressions other than neutral found in this image for calibration.

Consequently, the facial expressions are more balanced and personal biases towards a certain expression are removed. The effect can best be illustrated by an example. For instance, for a person a value of 0.3 for angry was found in the most neutral image. This means that for this test person 'angry' should be classified only when its value is higher than 0.3. The figure below shows how the classifier outputs are mapped to different values to negate the test person's bias towards 'angry'.

Continuous calibration continuously calculates the average expression of the test person. It uses that average to calibrate in the same manner as with the participant calibration.

An example of a possible classifier output correction for a specific facial expression using participant calibration.

——— uncalibrated
——— calibrated
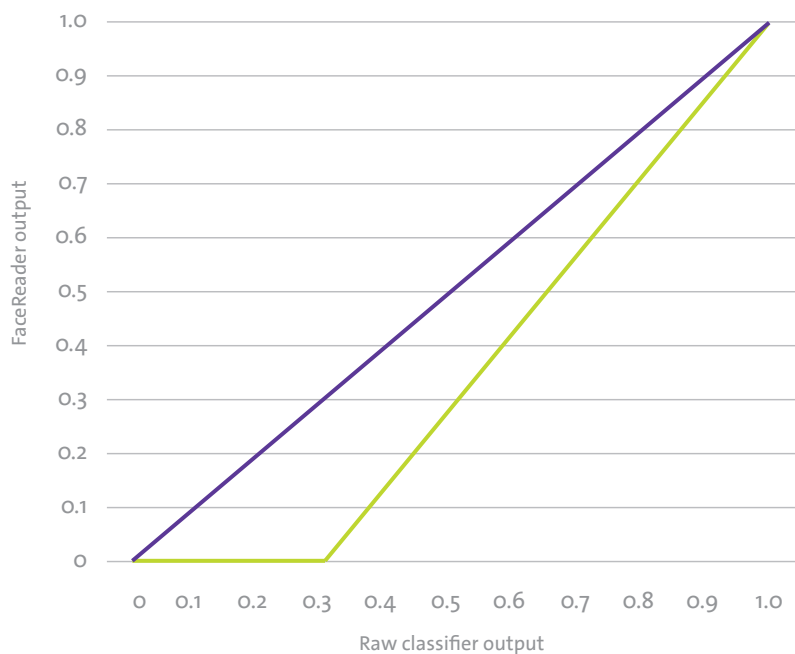
FACEREADER'S OUTPUT

FaceReader's main output is a classification of the facial expressions of your test participant. These results are visualized in several different charts and can be exported to log files. Each expression has a value between 0 and 1, indicating its intensity. '0' means that the expression is absent, '1' means that it is fully present. FaceReader has been trained using intensity values annotated by human experts.

Facial expressions are often caused by a mixture of emotions and it is very well possible that two (or even more) expressions occur simultaneously with a high intensity. The sum of the intensity values for the expressions at a particular point in time is, therefore, normally not equal to 1.

VALENCE

Besides the intensities of individual facial expressions FaceReader also calculates the valence. The valence indicates whether the emotional state of the subject is positive or negative. 'Happy' is the only positive expression, 'sad', 'angry', 'scared' and 'disgusted' are considered to be negative expressions. 'Surprised' can be either positive or negative and is, therefore, not used to calculate valence. The valence is calculated as the intensity of 'happy' minus the intensity of the negative expression with the highest intensity. For instance, if the intensity of 'happy' is 0.8 and the intensities of 'sad', 'angry', 'scared' and 'disgusted' are 0.2, 0.0, 0.3, and 0.2, respectively, then the valence is 0.8 – 0.3 = 0.5.

> The valence indicates whether the emotional state of the subject is positive or negative.

*Example of a Valence chart showing the valence (—) over time.*
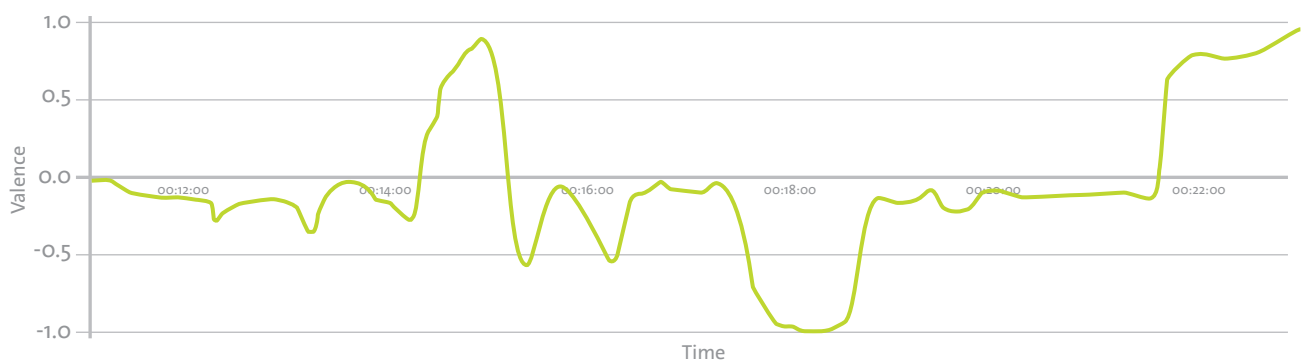
AROUSAL

Facereader also calculates arousal. It indicates whether the test participant is active (+1) or not active (0). Arousal is based on the activation of 20 Action Units (AUs) of the Facial Action Coding System (FACS) [10].

Arousal is calculated as follows:

1. The activation values (AV) of 20 AUs are taken as input. These are AU 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 20, 23, 24, 25, 26, 27, and the inverse of 43. The value of AU43 (eyes closed) is inverted because it indicates low arousal instead of high arousal like the other AUs.

2. The average AU activation values (AAV) are calculated over the last 60 seconds. During the first 60 seconds of the analysis, the AAV is calculated over the analysis up to that moment. *AAV = Mean (AVpast 60 seconds)*

3. The average AU activation values (AAV) are subtracted from the current AU activation values (AV). This is done to correct for AUs that are continuously activated and might indicate an individual bias. This results in the Corrected Activation Values (CAV). *CAV = Max(0, AV – AAV)*

4. The arousal is calculated from these CAV values by taking the mean of the five highest values. *Arousal = Mean (5 max values of CAV)*

Active

Unpleasant

Pleasant

HAPPY

Inactive

0%

5%

9%

14%

18%

23%
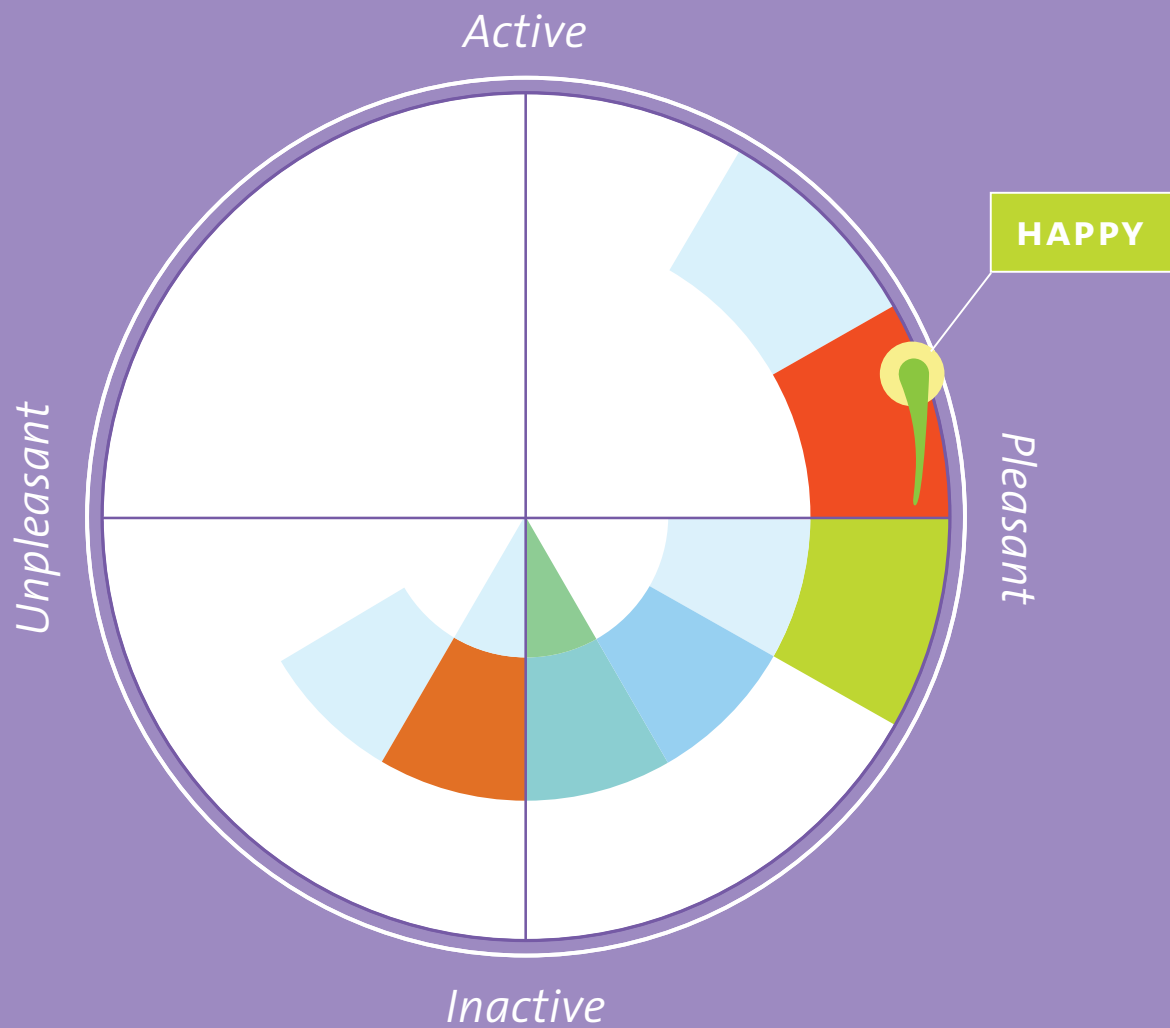
CIRCUMPLEX MODEL OF AFFECT

FaceReader's circumplex model of affect is based on the model described by Russel [11]. In the circumplex model of affect, the arousal is plotted against the valence. During the analysis, the current mix of expressions and Action Units is plotted with unpleasant/pleasant on the x-axis and active/inactive on the y-axis. A heatmap visualizes which of these expressions was present most often during the test.

ADD-ON MODULES

Several add-on modules expand FaceReader software to meet your research needs.

THE PROJECT ANALYSIS MODULE

With the Project Analysis module, an add-on module for FaceReader, you can analyze the facial expressions of a group of participants. You can create these groups manually, but you can also create groups based on the value of independent variables. By default the independent variables *Age* and *Gender* are present, which allows you to create groups with males and females, or age groups. You can also add independent variables to create groups. Add, for example, the independent variable *Previous experience* to create a group with participants that worked with a program before and a group with those that did not.

You can mark episodes of interest, for example the time when the partici-pants were looking at a certain video or image. This makes FaceReader a quick and easy tool to investigate the effect of a stimulus on a group of participants.

Simultaneous visualization of facial expressions of a group of participants, the stimulus video and the participant's face.
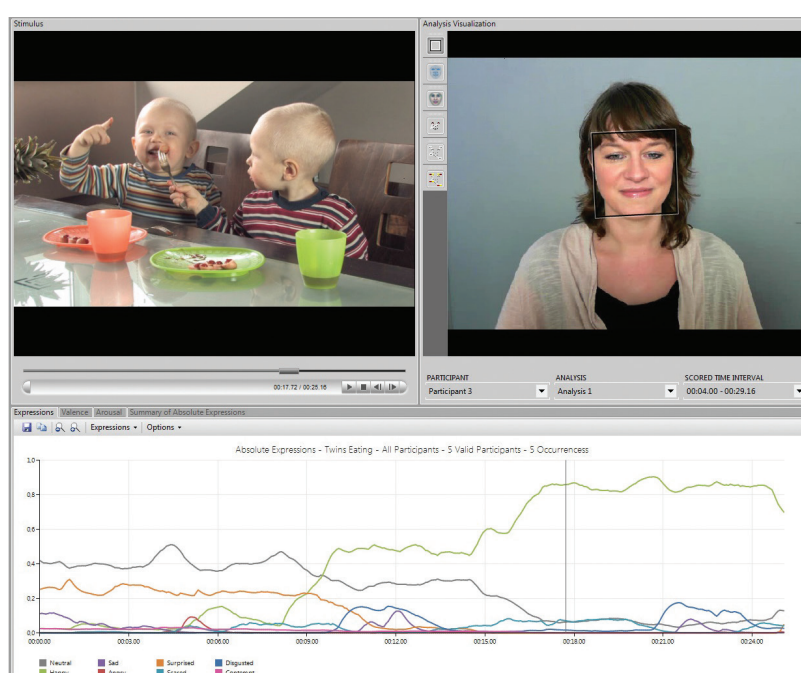
The numerical group analysis gives a numerical and graphical representation of the facial expressions, valence and arousal per participant group. With a click on a group name a T-test is carried out, to show in one view where the differences are.

The temporal group analysis shows the average expressions, valence and arousal of the group over time. You can watch this together with the stimulus video or image and the video of a test participant's face. This shows the effect of the stimulus on the participant's face in one view.

THE ACTION UNIT MODULE

Action Units are muscle groups in the face that are responsible for facial expressions. The Action Units are described in the Facial Action Coding System (FACS) that was published in 2002 by Ekman *et al*. [10].With the Action Unit Module, FaceReader can analyze 20 Action Units. Intensities are annotated by appending letters, A (trace); B (slight); C (pronounced); D (severe) or E (max), also according to Ekman *et al*. [10]. Export in detailed log as numerical values is also possible.

Action Unit classification can add valuable information to the facial expressions classified by FaceReader. The emotional state Confusion is, for example, correlated with the Action Units 4 (Brow lowerer) and 7 (Eyelid tightener) [12].

Action Unit classification can add valuable information to the facial expressions classified by FaceReader.

VALIDATION

To validate FaceReader, its results (version 7) have been compared with those of intended expressions. The figure on the next page shows the results of a comparison between the analysis in FaceReader and the intended expressions in images of the Amsterdam Dynamic Facial Expression Set (ADFES) [13]. The ADFES is a highly standardized set of pictures containing images of eight emotional expressions. The test persons in the images have been trained to pose a particular expression and the images have been labeled accordingly by the researchers. Subsequently, the images have been analyzed in FaceReader. As you can see, FaceReader classifies all 'happy' images as 'happy', giving an accuracy of 100% for this expression.

	NEUTRAL	HAPPY	SAD	ANGRY	SURPRISED	SCARED	DISGUSTED
NEUTRAL	95.5% (21)		4.3% (1)				
HAPPY		100% (23)					
SAD			91.3% (21)	4.0% (1)			
ANGRY				96.0% (24)			4.5% (1)
SURPRISED	4.5% (1)				100% (21)	4.3% (1)	
SCARED			4.3% (1)			95,6% (22)	4.5% (1)
DISGUSTED							90.9% (20)
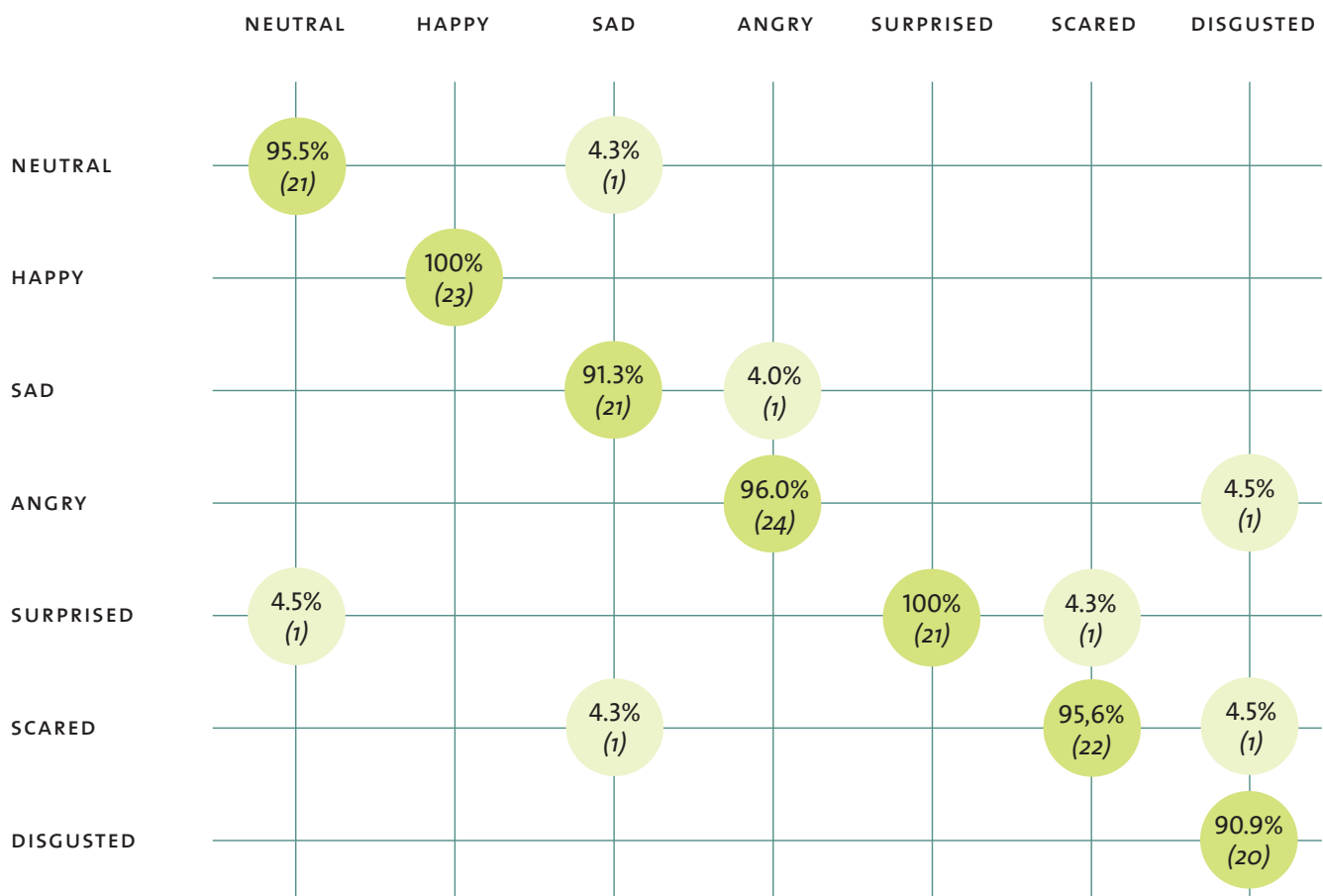
VALIDATION OF ACTION UNIT CLASSIFICATION

The classification of Action Units has been validated with a selection of images from the Amsterdam Dynamic Facial Expression Set (ADFES) [13] that consists of 23 models performing nine different emotional expressions (anger, disgust, fear, joy, sadness, surprise, contempt, pride, and embarrassment). FaceReader's classification was compared with manual annotation by two certified FACS coders. For a detailed overview of the validation, see the paper Validation Action Unit Module [14] that you can obtain from your Noldus sales representative.

POSED OR GENUINE

Sometimes the question is asked how relevant the results from FaceReader are if the program has been trained using a mixture of intended and genuine facial expressions. It is known that facial expressions can be different when they are intended or genuine. An intended smile is for example characterized by lifting the muscles of the mouth only, while with a genuine smile the eye muscles are also contracted [15]. On the other hand, one could ask what exactly a genuine facial expression is. Persons watching a shocking episode in a movie may show very little facial expressions when they watch it alone. However, they may show much clearer facial expressions when they watch the same movie together with others and interact with them. And children that hurt themselves often only start crying once they are picked up and comforted by a parent. Are those facial expressions that only appear in a social setting intended or genuine? Or is the question whether a facial expression is genuine or intended perhaps not so relevant?

FaceReader does not make a distinction whether a facial expression is acted or felt, authentic or posed. There is a very high agreement with facial expressions perceived by manual annotators and those measured by FaceReader [13]. One could simply say that if we humans experience a face as being happy, FaceReader detects it as being happy as well, irrespective from whether this expression was acted or not.

ARE FACIAL EXPRESSIONS ALWAYS THE SAME?

Another frequently asked question is whether the facial expressions measured by FaceReader are universal throughout ages, gender, and culture. There are arguments to say yes and to say no. The fact that many of our facial expressions are also found in monkeys supports the theory that expressions are old and therefore are independent of culture. In addition to this, we humans have until not so long ago largely been unaware of our own facial expression, because we did not commonly have access to mirrors. This means that facial expressions cannot be explained by copying behavior.

Furthermore, people that are born blind have facial expressions that resemble those of family members. This indicates that these expressions are more likely to be inherited than learned.

On the other hand, nobody will deny that there are cultural differences in facial expressions. For this purpose, FaceReader has different models, for example the East-Asian model. These models are trained with images from people of these ethnic groups. And it is true that with the East Asian model FaceReader gives a better analysis of facial expressions of East Asian people than with the general model and vice versa. But this effect is very small, there is only a 1 to 2 percent difference in classification error. These are all arguments supporting the statement made by Ekman & Friesen [1] that the seven facial expressions are universal and can reliably be measured in different cultures.

Feel free to contact us or one of our local representatives for more references, clients lists, or more detailed information about FaceReader and The Observer XT.

www.noldus.com

REFERENCES

1   P. Ekman (1970). Universal facial expressions of emotion. *California Mental Health Research Digest*, **8**, 151-158.

2   P. Ekman and W. V. Friesen (1986). A new pan-cultural facial expression of emotion. *Motivation and emotion*, **10(2)**, 159-168.

3   H. van Kuilenburg, M. Wiering and M.J. den Uyl (2005). Model Based Method for Automatic Facial Expression Recognition. *Proceedings of the 16th European Conference on Machine Learning, Porto, Portugal, 2005, pp. 194-205, Springer-Verlag GmbH*.

4   M.J. den Uyl and H. van Kuilenburg (2008). The FaceReader: Online Facial Expression Recognition. *Proceedings of Measuring Behavior 2005, Wageningen, The Netherlands, August 30 - September 2, 2008*, pp. 589-590.

5   H. van Kuilenburg, M.J. den Uyl, M.L. Israël and P. Ivan (2008). Advances in face and gesture analysis. *Proceedings of Measuring Behavior 2008, Maastricht, The Netherlands, August 26-29, 2008*, pp. 371-372.

6   P. Viola and M. Jones. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, U.S.A., December 8-14, 2001*.

7   T. Cootes and C. Taylor (2000). *Statistical models of appearance for computer vision*. Technical report, University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering.

8   C.M. Bishop (1995). *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford.

9   A. Gudi, H.E. Tasli, T. M.den Uyl and A. Maroulis (2015). *Deep Learning based FACS Action Unit Occurrence and Intensity Estimation. Automatic Faces and Gesture Recognition (FG)*.

10 P. W. Ekman, V. Friesen and J. C. Hager (2002). *FACS manual. A Human Face*.

11 J. Russell, (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, **39**, 1161–1178.

12 J. F. Grafsgaard, J. B. Wiggins, K. E. Boyer, E. N. Wiebe and J. C. Lester (2013). Automatically recognizing facial expression: Predicting engagement and frustration. In *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 43-50).

13 J. van der Schalk, S. T. Hawk, A. H. Fischer and B. J. Doosje (2011). Moving faces, looking places: The Amsterdam Dynamic Facial Expressions Set (AD-FES). *Emotion*, **11**, 907-920. DOI: 10.1037/a0023853

14 P. Ivan and A. Gudi (2016). *Validation Action Unit Module*.

15 P. Ekman and W. Friesen (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior*, **6(4)**, 238-252.

N