Supporting Information

Du et al. 10.1073/pnas.1322355111

SI Text

Subordinate Relationship. In Fig. 2, we showed six examples of the subordinate relationship between six compound facial expressions of emotion and their corresponding basic categories. In Fig. S1, we detail the subordinate relationships for the remaining nine compound categories. As in Fig. 2, this figure illustrates which prototypical action units (AUs) are kept and which need to be eliminated when producing a facial expression of a compound emotion category. The AUs that are eliminated cannot be kept because they cannot be readily produced simultaneously with the others. As a result, some emotion categories are easier to recognize than others; this was evident in the results of Table 4. These coarticulations and how they are generated are essential for disambiguating expressions such as appalled and hate, which, despite of having the same subordinate basic emotion categories, are produced quite distinctly, which facilitates their recognition even if the intensities of their AUs are not known.

In Table 1, we also noted that some less common AUs are added to the production of a facial expression of a compound emotion; i.e., some facial expressions of compound emotion categories include AUs that are not present in either of their subordinate categories. In some cases, an additional AU may be elicited to facilitate the recognition of the compound expression. For example, sadly disgusted includes AU 25; this emphasizes disgust while keeping the corners of the mouth pointing downward to specify sadness. Fig. S1 also shows a few AU particularities of this subject. The AUs identified in gray font are those used by this subject but not generally observed in the majority. The subject shown in this figure, for instance, has a preference for incorporating AU 11, which creates the wrinkle in the nasolabial area.

Database. ***Participants.*** Subjects were asked to practice each expression before photos were taken. A mirror was placed next to the camera to facilitate practicing while looking forward. During acquisition, the experimenter instructed the subject to look frontally at the camera. This allowed us to obtain fully frontal views of the subjects' faces. Subjects were given the opportunity to take breaks to avoid production differences between facial expressions due to boredom or muscle fatigue. The picture of the production was taken at the apex of the expression. Before proceeding, this picture was visually evaluated by the experimenter, e.g., to guarantee that the picture was taken at the apex and that the subject was facing frontally. If the experimenter believed one of these problems occurred, the subject was asked to repeat that expression.

Participants were non-FACS coders and were naive to the scientific question or procedure used to analyze the images. Subjects were asked to think of an experience where the facial expression of emotion to be produced was typical. Sample scenarios were given to participants to facilitate this process. Subjects were not instructed to move any specific facial muscles. Rather, they were asked to express each emotion category as clearly as possible according to the provided definitions and the sample situations where such an expression would be produced.

Detection errors. The results in Table 2 are the average pixel differences between the specified computational models and the manual markings obtained by the authors. Manual markings (as given by a trained eye) have proven invaluable to test face and facial feature detection algorithms. We manually marked the coordinates of 94 fiducial points (i.e., face landmarks) in all of the 5,060 images in our database. These landmarks were carefully selected to represent the contour of the face (i.e., jaw line) as well as the shape of each of the internal facial components (Fig. 3). Manually marking the same landmark in each face would of course be impossible to do by a human—each time marking a close-by pixel, not the intended one; this is resolved as follows. Once the 94 fiducials are marked in an image, the best-fitted cubic splines for each of the seven facial components (eyes, eyebrows, nose, mouth, face outline) are computed. Each spline is then resampled by a constant number of equally distant landmarks. We used 8 fiducials to represent each brow, 9 for each eye, 15 for the nose, 30 for the mouth, and 15 for the face outline. The same spline and resampling procedure was used for each of the detection algorithms. The detection errors in Table 2 were then computed as follows. For each of the algorithms used, the average Euclidean distance between the manual and automatic detected landmarks defining each facial component (eyes, eyebrows, nose, mouth, face outline) is computed. Then the average of all of the landmark points is calculated. These Euclidean distances are given in image pixels.

Fig. S1. Shown here are the AUs of the remaining nine facial expressions of compound emotion categories. The prototypical AUs of the basic emotions are combined as shown to produce the compound category. The prototypical and variant AUs of the basic expressions kept to produce the compound emotion are marked with a bounding box. These relationships define the subordinate classes of each category. The results thus define possible confusion of the compound emotion categories by their subordinates and vice versa. AUs only used by these subjects (or a small number of subjects, smaller than 20%) are shown in gray font; these represent particularities of the subject.