© Académie des sciences / Elsevier, Paris

A two-process theory of facial perception of emotions

La perception des expressions faciales émotionnelles. Une théorie à deux processus

Monique de Bonis[a]*, Paul De Boeck[b], Fernando Pérez-Diaz[c], Monique Nahas[d]

[a]CNRS-UMR 7593, faculté de médecine Paris-Sud, 63, rue Gabriel-Péri, 94276 Le Kremlin-Bicêtre, France
[b]CNRS-UMR 7593, hôpital de la Salpêtriere, 47, boulevard de l'Hôpital, 75013 Paris, France
[c]Department of Psychology, Catholic University of Leuven, Tiensestraat 102, 3000 Leuven, Belgium
[d]Département « Image numérique », université Paris-VIII, 2, rue de la Liberté, 93526 Saint-Denis, France

(Received 8 March 1999; accepted 13 April 1999)

Note communicated by Pierre Buser

Abstract — A two-process probabilistic theory of emotion perception based on a non-linear combination of facial features is presented. Assuming that the upper and the lower part of the face function as the building blocks at the basis of emotion perception, an empirical test is provided with fear and happiness as target emotions. Subjects were presented with prototypical fearful and happy faces and with computer-generated chimerical expressions that were a combination of happy and fearful. Subjects were asked to indicate the emotions they perceive using an extensive list of emotions. We show that some emotions require a conjunction of the two halves of a face to be perceived, whereas for some other emotions only one half is sufficient. We demonstrate that chimerical faces give rise to the perception of genuine emotions. The findings provide evidence that different combinations of the two halves of a fearful and a happy face, either congruent or not, do generate the perception of emotions other than fear and happiness. © Académie des Sciences / Elsevier, Paris

facial expression of emotion / perception / probabilistic model of classification

Résumé — Dans cet article on propose une théorie de nature probabiliste, fondée sur deux processus et destinée à rendre compte de la perception des émotions exprimées par un visage. Partant de l'hypothèse selon laquelle les parties supérieure et inférieure d'un visage constituent les unités de base sur lesquelles se fonde la perception d'une émotion, une vérification empirique de la théorie est apportée avec la peur et la joie comme émotions cibles. Le protocole expérimental a consisté à présenter à des sujets des visages exprimant des émotions prototypiques de joie et de peur ainsi que des visages chimériques construits par ordinateur et composés par la fusion de parties supérieure et inférieure des deux expressions de peur et de joie, Les sujets, étudiants en médecine de langue française, devaient déterminer la ou les émotions exprimées par chaque visage à partir d'une liste extensive d'émotions. On a montré que certaines émotions réclament la conjonction des deux parties du visage alors que pour d'autres, une seule partie est suffisante; que les expressions chimériques donnent lieu à la

* Correspondence and reprints: Monique.Debonis@kb.u-psud.fr

perception d'émotions authentiques. Les données apportent la preuve que différentes combinaisons des deux parties d'une expression de peur ou de joie, concordantes ou non, génèrent la perception d'émotions autres que la peur ou la joie, © Académie des Sciences / Elsevier, Paris

expression faciale des émotions / perception / modèle de classification probabiliste

Version abrégée

Dans la présente étude on s'est interrogé sur la nature des mécanismes qui règlent, chez le sujet humain, la perception des émotions exprimées par un visage. Deux questions ont été examinées : existe-t-il des unités faciales expressives primitives, et, si oui, de quelle façon ces primitives se combinent-elles pour déterminer la perception d'une catégorie particulière d'émotion et pour la différencier d'une catégorie voisine ou éloignée dans la structure du système des émotions. Nous avons proposé un modèle de classification hiérarchique de type probabiliste pour rendre compte de ces mécanismes perceptifs. La propriété principale de ce modèle est de définir l'appartenance d'un élément (i.e. une expression faciale, ou une partie de cette expression) à une catégorie (i.e. une émotion) en termes de processus stochastique et non en tout ou rien. Considérant le haut et le bas d'un visage comme des unités primitives, nous avons appliqué ce modèle à la perception de visages prototypiques de joie et de peur et à des visages chimériques. Ces visages chimériques, construits par ordinateur, ont été obtenus par la fusion, selon un programme d'image numérique, de la partie inférieure et supérieure des deux émotions prototypiques. Le protocole expérimental a consisté à recueillir chez des sujets les jugements perceptifs de ces visages sur la base d'une liste extensive d'émotions. Les résultats ont montré que ce modèle probabiliste rendait compte de façon satisfaisante des données expérimentales. Nous avons démontré que certaines émotions étaient perçues sur la base d'une seule partie du visage (émotions à un seul trait), alors que d'autres émotions nécessitaient la conjonction des deux parties (émotions à double trait). De plus, il est apparu que les expressions chimériques, loin de donner lieu à des jugements perceptifs incongrus, engendraient la perception d'émotions authentiques, qualitativement différentes des émotions de base. Il existe donc de fortes présomptions pour que des combinaisons différentes des deux parties d'un visage exprimant deux émotions contrastées génèrent la perception d'émotions émergentes. La généralisation du modèle devrait permettre d'isoler les primitives expressives sur la base desquelles repose la perception des émotions.

1. Introduction

The nature of the processes underlying facial expressions of emotions was examined. To explain how facial expressions lead to the perception of emotions, a two-process probabilistic theory was postulated [1]. The first process is the activation of features from the perception of a facial expression, and the second is the association of an emotion to the activated pattern of features. The association of features to emotions is considered to be conjunctive. Both feature activation and their conjunctive association to emotions are assumed to be probabilistic. It is hypothesised that the features correspond to separate configurations of the upper part and the lower part of the face, which combine into one global facial expression.

Regarding the feature activation process, it is assumed that while a face is perceived, features are activated with a certain probability, called feature activation probability. The result of this process is a pattern of activated features. Regarding the feature–emotion association process, two assumptions are made. First, each emotion is assumed to set conditions for being perceived, with the conditions referring to features that must be activated. Whether a feature actually is a condition in a given situation, is not an all-or-none issue, but a stochastic characteristic of the emotion. Emotions are supposed to be characterised by a set of feature requirement probabilities, called feature–emotion association probabilities. Second, for an emotion to be perceived, all features that are actually required on the basis of the stochastic process need to be activated while perceiving a face. That all required features need to be activated is what makes the feature–emotion association process conjunctive. This quality gives the perception of emotions in faces a Gestalt-like character [2, 3]. From the stochastic conjunction of features associated with an emotion, it can be derived what the probabilities are for each pattern of activated features. The latter probabilities are called the conjunctive association probabilities to be distinguished from the more elementary feature association probabilities. Given that the feature–emotion association is conjunctive, two types of emotions can be discerned: single-feature emotions and multiple-feature emotions. For the former, an association with one feature is a sufficient condition, as in a disjunctive process; for the latter, more than one feature is required and they all need to be activated. We hypothesise that the upper part and the lower part of the face function as building blocks at the basis of emotion perception. We do not imply, however, that these building blocks cannot be further analysed into smaller units as in several descriptive taxonomies of facial expression [4–6]. The only claim is that it makes sense from the neuropsychology of facial expression of emotion

to study facial emotion perception at the level of upper and lower face-halves [7]. In other words, within-half configurations can be seen as features that are associated with the perception of emotions, albeit that they are rather global features. According to this face-half view, single-feature emotions are emotions that are associated with only one half of the face, either its upper or its lower part, and, when an emotion is associated with a feature conjunction, then it corresponds to a combined configuration, one in the upper part and another in the lower part. Although there has been some research on the contribution of the upper part and lower part of the face [8, 9] thus far it has not been investigated systematically what the role of each part is. Morphing techniques allow us to create combinations of upper and lower parts to test our face-half hypothesis. Starting from faces with clearly identified emotions, the upper and lower parts of a facial expression were, according to a new numerical image algorithm, continuously fused into a realistic but 'chimerical' expression [10].

The aim of the present study was to provide an empirical test of the theory with fear and happiness as target emotions. The reason for the choice of these two emotional expressions is that the neural bases of their differentiation have been clearly identified [11].

2. Materials and methods

Four types of faces were considered:

– HH faces, with an upper and a lower happy part;

– FF faces, with a fearful upper and lower part.

– HF faces, with an upper happy part and a fearful lower part;

– FH faces, with a fearful upper part and a happy lower part.

The visual material is composed of a set of 20 photographed faces, four faces of each of five persons (two males and three females).

For each person, one face was a HH face and one other a FF face. These ten faces were taken from a standard set of pictures of emotional facial expressions [12]. For each person, also two chimerical faces (HF, FH) were generated by a computer-based numerical image manipulation of the prototype of each emotional expression, either fear or happiness (*figure 1*). Since the procedure used to create the chimerical faces is new, a detailed description is given below.

2.1. The numerical image procedure to generate chimerical facial expressions

An interpolation between two pictures can be obtained from common morphing techniques. The principle is to define two sets of corresponding segments in the initial and final picture. Each couple of segments is interpolated linearly using a parameter t in the range 0–1. One can use a segment to build a co-ordinate referential, so that any given point in the interpolated picture has two co-ordinates relative to the interpolated segment; the same co-ordinates are used to map the point on the start and end pictures with the original segments. The interpolated pixel value is a mixture of the mapped pixel values. In order to use all the segments simultaneously, we mixed the results for all independent segments according to the distance from the interpolated point to the interpolated segments. In the present case, we used two sets of segments and mixed them with a parameter which is not constant, as in previous studies, but which was varying within spatial co-ordinates, in particular only along the vertical axis. The top of the first picture can be continuously mixed with the bottom of the second picture. The parameter's fastest variation was set to be at the middle of the nose where the interpolation of the skin is not crucial. As compared to the morphing procedures used in previous studies [13, 14], the major difference is the variation of the morphing coefficient within spatial co-ordinates.

2.2. Subjects

Eighteen right-handed healthy volunteers, French medical students at University Paris-XI (eight males and ten females, mean age 22.9 years, std 3.62), were presented with the 20 faces, each on one black-and-white slide, arranged in a random order, and presented on a screen for 1 min. The subjects were asked to indicate as many emotions as they perceived in each face, using an extended list of 19 emotion-words, taken from a study by Rosenberg and De Boeck [15] and translated into French. This list includes not only words denoting the five basic emotions: happiness, fear, sadness, anger and disgust, plus surprise, but also words sampled from an extended lexicon of emotions [16, 17].

Nine words are positive: admiration, affection, amused, cheerful, connected, enjoyment, interested, relaxed, warm; nine are negative: angry, confused, contempt, distant, embarrassed, fearful, pained, repulsed, sad and one is neutral: surprise.

The agreement among the 18 subjects assessed by the intraclass coefficient ICC (2k) [18] was 0.82 with a 95 % confidence interval of 0.74–0.90.

3. Results

The resulting 20 faces × 19 emotion-word data matrix, with frequencies of perceived emotions in the cells was processed with a feature extraction algorithm that is directly based on the assumed formal feature-based process of a conjunctive nature [19].

The algorithm is a maximum likelihood algorithm. The analysis consists in extracting a number of hypothetical features and probabilities related to these features, such that the likelihood of the frequencies is maximised. For a good understanding of the results, a brief description of the mathematical model used in processing the data is given below.

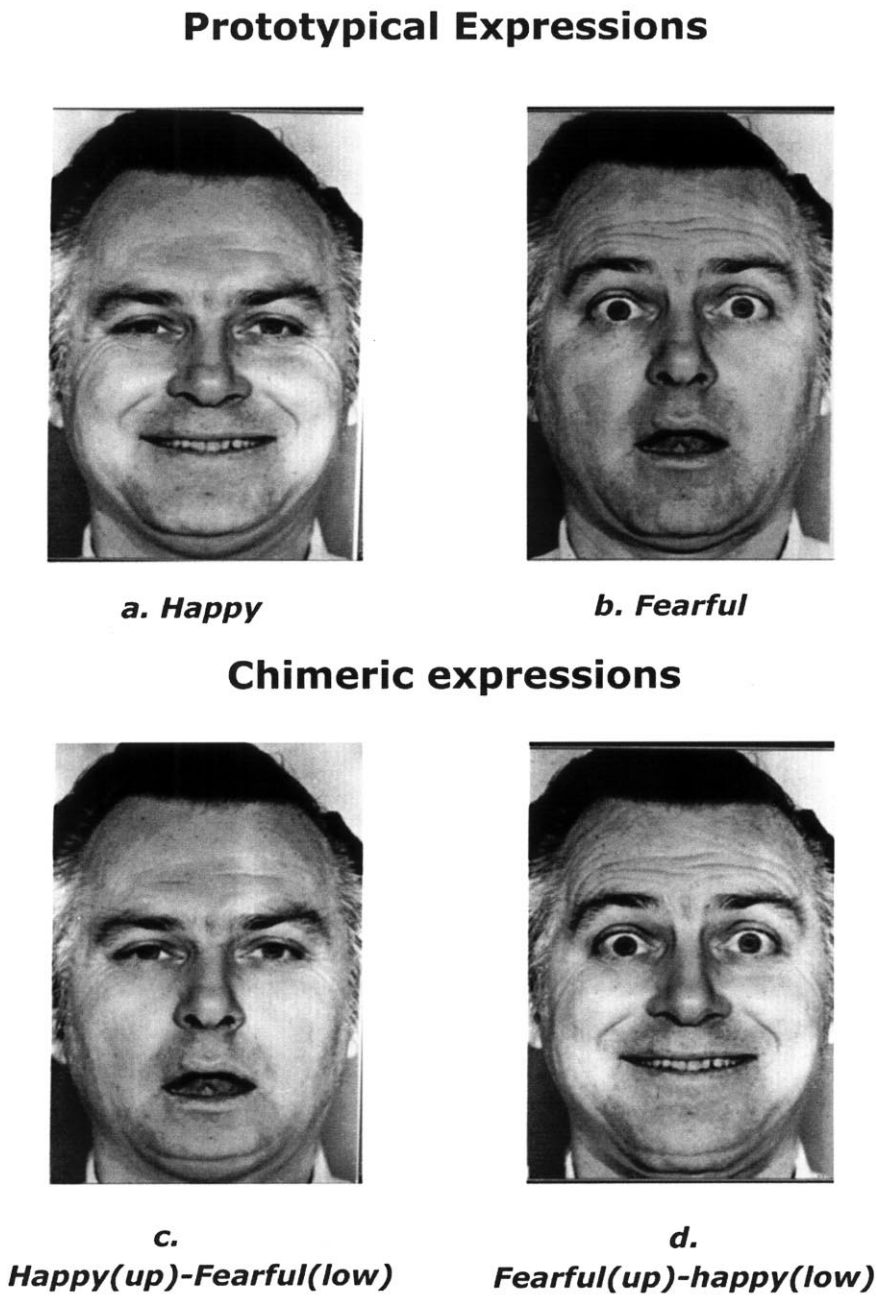


Figure 1. One example of the pictures used in the experiment.
Faces (a) and (b) are prototypical happy and fearful expressions of a single individual from a standard set. Faces (c) and (d) are computer-generated expressions in which the upper and the lower parts of the two prototypical expressions have been fused to create a chimerical expression. Face (c) is made of the happy upper part of (a) and the fearful lower part of (b); face (d) is made of the fearful upper part of (b) and the happy lower part of (a).

3.1. The probabilistic model for non-linear combination of features

The underlying model, called the probability matrix decomposition model (PMD), is a probabilistic version of the HICLAS model [20]. HICLAS can be considered as a two-way simultaneous classification model. It is a classification model indeed since an extracted feature can be seen as a classification variable, and it is a two-way simultaneous classification model since the extracted features are simultaneously linked to row and column elements, faces and emotions in this case. The probabilistic nature of PMD adds a probabilistic aspect to the classification, meaning that belonging to a category is not all-or-none, but a stochastic event instead. The hierarchical aspect of the model is reflected in that the entities, in this case faces and emotions, can be described as sets of features, so that one can map superset and subset relations between these sets. The estimation method is based on an

Table I. Mean feature activation probabilities per type of face.

	Hypothetical features			
Faces	1 FL	2 FU	3 HL	4 HU
HH	0.067	0.044	0.976	0.881
FF	0.829	0.899	0.046	0.286
HF	0.857	0.363	0.245	0.690
FH	0.205	0.912	0.768	0.241

H: happy; F: fearful; U: upper; L: lower.

EM algorithm complemented with a Bayesian prior, so that the parameter estimates are maximum a posteriori (MAP) values.

Two kinds of probabilities are estimated in the analysis, these probabilities are the parameters of the underlying model. The first kind consists of feature probabilities of the faces, the so-called feature activation probabilities. The second consists of feature probabilities associated with the emotions, indicating the probability that an emotion requires the corresponding feature to be activated for the emotion to be perceived. From these two kinds of probabilities one can derive the conjunctive association probabilities, that is, of perceived emotions being associated with a pattern of activated features.

3.2. Feature activation probabilities

The conjunctive analysis was run from 1 to 6 features. From 1 to 6 features, the correlation between the observed data and the data predicted by the model was: 0.452, 0.905, 0.958, 0.969, 0.971 and 0.974. The goodness of fit did not much improve from more than four features: in terms of proportion a mean absolute difference between observed and expected frequencies of 0.036, thus the four features were retained as the optimal solution.

These four hypothetical features as derived from the analysis were completely in line with the way the stimuli have been designed. The average feature probabilities for the four types of faces are shown in *table I*.

The four features are clearly linked to the four types of halves. The first feature is linked to F lower, the second to F upper, the third to H lower, and the fourth to H upper. For a good understanding, it should be stressed that the statistical analysis is completely blind with respect to the design of the stimulus set. It was a purely empirical result that four hypothetical features were discovered in the data, and that they turned out to be these four.

3.3. Feature probabilities of the emotions

Regarding the emotions two kinds of probabilities can be reported. We will first report the feature association probabilities, as they are parameters of the model (*table II*) and then the conjunctive association probabilities which are derived from the former. The following hypothetical example explains how the conjunctive association probabilities can be derived from the feature association probabilities associated with the emotion. Suppose a given face

Table II. Feature–emotion association probabilities.

Emotions	1 FL	2 FU	3 HL	4 HU
Admiration	0.179	0.608	0.881	0.337
Affection	0.030	0.030	0.959	0.504
Amused	0.038	0.082	0.980	0.075
Cheerful	0.013	0.013	0.972	0.051
Connected	0.030	0.035	0.546	0.150
Enjoyment	0.021	0.022	0.969	0.151
Interested	0.062	0.118	0.498	0.165
Relaxed	0.019	0.022	0.939	0.502
Warm	0.024	0.026	0.947	0.219
Angry	0.987	0.328	0.352	0.971
Confused	0.651	0.951	0.076	0.547
Contempt	0.973	0.340	0.277	0.941
Distant	0.954	0.109	0.079	0.403
Embarrassed	0.635	0.946	0.060	0.077
Fearful	0.915	0.953	0.035	0.034
Pained	0.979	0.793	0.350	0.226
Repulsed	0.969	0.631	0.057	0.220
Sad	0.977	0.895	0.352	0.265
Surprise	0.040	0.978	0.093	0.024

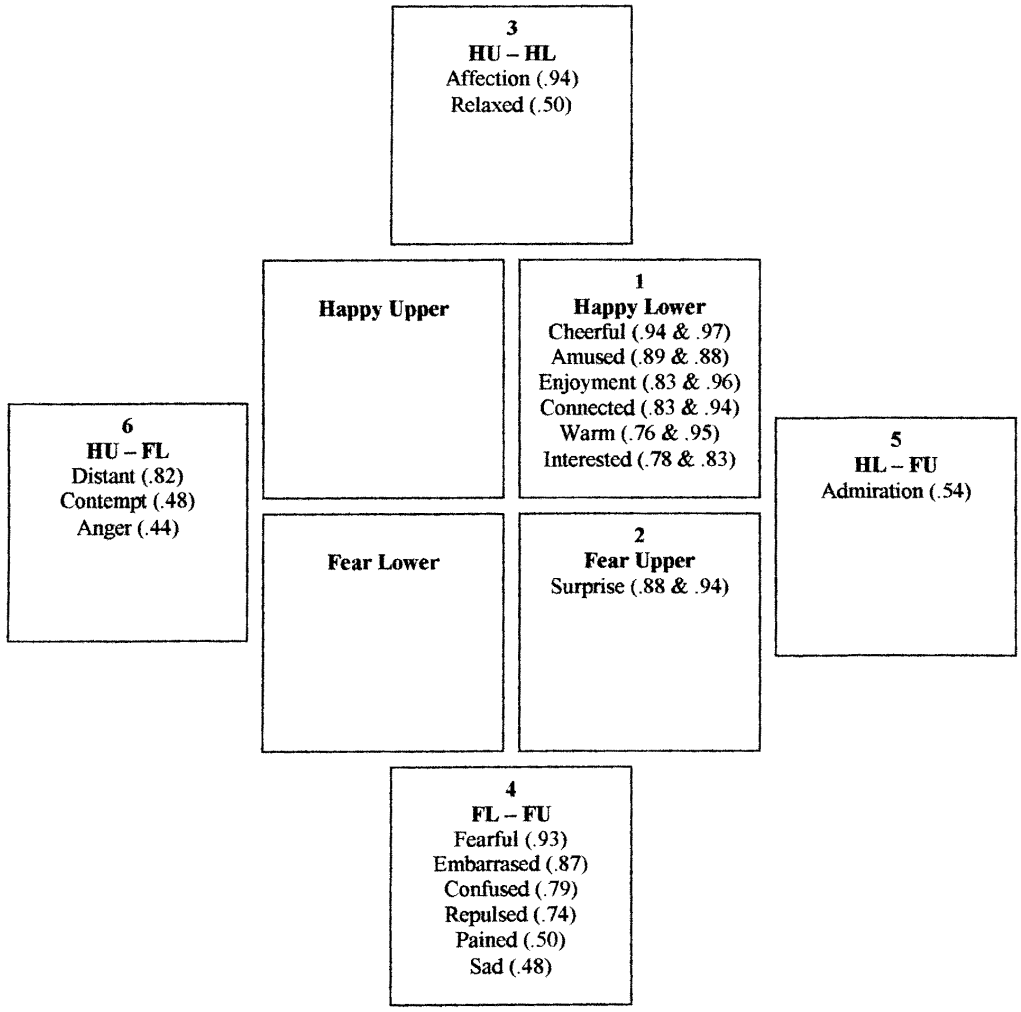
FL: fear lower; FU: fear upper; HL: happy lower; HU: happy upper.

actually activates two features out of four, namely the second and the third, which can be represented as (0, 1, 1, 0), and that the feature–emotion association probabilities of an emotion are as follows: 0.70, 0.20, 0.80, 0.30. This specific set of probabilities means that in most cases the emotion in question requires features 2 and 3 to be activated. The probability for the face in question to be perceived as expressing the emotion in question is $(1.00 - 0.70) \times (1.00) \times (1.00) \times (1.00 - 0.30) = 0.21$. Given that the face does activate features 2 and 3, the emotion conditions related to these features are certainly met. Given that features 1 and 4 are not activated, the emotion can be perceived only if the emotion does not require these features to be activated, the probability of this per feature being the complement of the probability that activation is required, that is $(1.00 - 0.70)$ and $(1.00 - 0.30)$, respectively.

The joint structure derived from the formal feature-based model is shown in *figure 2*. It can be derived from the four patterns of activated features whether the perception of an emotion is based on one single feature or on a combination of two features.

3.4. Single-feature emotions

Six positive emotion-words and surprise were categorised as single-feature emotions. For those positive emotions, feature 3, corresponding to the lower part of a happy face, is a sufficient feature, the upper part, either happy or fearful, does not matter. The best instances are 'amused' and 'cheerful', possibly expressing basic happiness, whereas most of the other emotion-words have a more complex relational meaning such as 'connected' and 'interested'. Similarly, feature 2, corresponding to fearful upper part, is a sufficient feature for 'surprise' to be perceived. Neither feature 1 nor feature 4 corresponding to the happy



H: Happy, F: Fearful, U: Upper, L: Lower

The numbers next to the emotion-words correspond to the conjunctive association probabilities. For single-Feature emotions, (central boxes 1 and 2) two probabilities are given. For example, for cheerful the left .94 correspond to the associated probability of a happy lower and upper facial pattern, and .the right .97 of a happy lower and a fearful upper facial pattern, both patterns share the happy lower part. For double-feature emotions (peripheral boxes 3,4,5 and 6) only one probability is given as they are associated primarily to only one pattern of features.

Figure 2. Structure of hypothetical facial patterns and their conjunctive association probabilities.

upper part and the fearful lower part, respectively, are sufficient to generate the perception of an emotion (empty boxes). It is noteworthy that none of the negative emotion-words belongs to the seven single-feature emotions.

3.5. Double-feature emotions

The remaining 12 emotion-words are allocated to the four boxes, two for positive and two for negative emotions. The basic emotion of fear requires both parts of a fearful face to be perceived. However, other negative emotions denoted by the words 'sad', 'repulsed', 'pain', 'embarrassed' could also be perceived, when the whole fearful facial pattern was activated, although the associated probabilities are not as high as for fear. Both the upper and the lower parts of a happy face are required for the perception of 'affection' and 'relaxed'. For example, for 'affection' the probability of being perceived is clearly higher when the features corresponding to the happy upper and happy lower are activated, as by HH faces, than for any other activation pattern.

The conjunction of contrasting parts of happy and fearful faces (chimerical expressions) gave rise to the perception of meaningful emergent emotions which are not perceived at all in the prototypical faces used to create the chimerical faces. Admiration (box 5) is perceived when the upper part shows fear, while the lower part shows

happiness; 'distant', 'contempt' and 'anger' are perceived when the upper part shows happiness, while the lower part expresses fear (box 6). The mixture of a happy and a fearful part never generates the perception of mixed-valence emotions.

4. Conclusion

Our two-process theory of emotion perception is supported by the data. We provide evidence that emotion perception is largely conjunctive, and it follows a probabilistic rule. Our face-half view allows a clear-cut distinction between single-feature and double-feature emotions. Basic positive emotions and surprise, seem to be perceived on the basis of only one part of an expressive face: the happy lower part and the fearful upper part, respectively. To the extent that each of these features are sufficient to produce the perception of an emotion, they could be considered as primitive facial units. All the other emotions investigated in this study require a combination of the two parts of an expressive face, either congruent or not. The basic emotion of fear was found to be a double-feature emotion. We show that basic negative emotions other than fear, such as sadness, pain and disgust could, with certain specified probabilities, be also perceived in response to the whole fearful facial pattern. Many emotions other than those related to fear and happiness can be understood from the different combinations of the upper and the lower part of a face. Happiness and fear seem to share facial ingredients with other emotions. A point of major interest is that the artificial expressions, the so-called chimerical faces, were perceived, with a strong between-subject agreement, as genuine, mainly complex emotions: admiration and contempt are cases in point. On the basis of our data we expect that facial mixtures, other than fear and happiness, created from the two halves of prototypic expressions of emotions, are primitive expressive units that could generate the perception of an even larger set of emotions than the one investigated. Such a result will provide a generalisation of the processes of emotion perception demonstrated here.

Acknowledgments: We thank P. Ekman for giving us permission to use and manipulate photographs from the Ekman and Friesen Picture of Facial Affect.

References

[1] Maris E., De Boeck P., Van Mechelen, I., Probability matrix decomposition models, Psychometrika 61 (1996) 7–29.

[2] Sergent N., An investigation into component and configural processes underlying face perception, Br. J. Psychol. 75 (1984) 221–242.

[3] Calder A.J., Young A.W., Perrett D.L., Etcoff N.L., Rowland D., Categorical perception of morphed facial expressions, Vis. Cognit. 3 (1996) 81–117.

[4] Duchenne De Boulogne G.B., Mécanismes de la physionomie humaine ou analyse électrophysiologique de l'expression des passions, J. Renouard, Paris, 1862.

[5] Darwin C., The Expression of Emotions in Man and Animals, J. Murray, London, 1872.

[6] Ekman P., Emotion and the Human Face, Cambridge Univ. Press, Cambridge, 1982.

[7] Berges J., Composantes tonico-motrices de la mimique, Entret. Pychiatr. 9 (1960) 41–69.

[8] Bassili J.N., Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face, J. Pers. Soc. Psychol. 37 (1979) 2049–2058.

[9] Katsikitis M., The classification of facial expressions of emotion: A multidimensional-scaling approach, Perception 26 (1997) 613–626.

[10] Nahas M., Huitric H., Computer program, Laboratoire d'image numérique, Univ. Paris-VIII, France, 1997.

[11] Morris J.S., Frith C.D., Perrett D.I., Rowland D., Young A.W., Calder A.J., Dolan R.J., A differential neural response in the human amygdala to fearful and happy facial expressions, Nature 383 (1996) 812–815.

[12] Ekman P., Friesen W.V., Facial Action Coding System, Palo Alto, Consulting Psychologists Press, 1976.

[13] Beier T., Shawn N., Feature-based image metamorphosis, Comput. Graph. 26 (1992) 35–42.

[14] Benson P.J., Perrett D.J., Synthesising continuous-tone caricatures, Image Vis. Comput. 9 (1991) 123–129.

[15] Rosenberg S., De Boeck P., Emotion depicted and experienced: Picasso's portraiture, in: Dorfman L., Martindale C., Leontiev D.G., Cupchick V., Petrov P., Matchotka N. (Eds.), Emotion, Creativity and Arts, Perm Institute of Art and Culture, Perm, 142, vol. 1997, pp. 371–386.

[16] Johnson-Laird P.N., Oatley K., The language of emotion: An analysis of the semantic field, Cogn. Emoti. 3 (1989) 81–123.

[17] de Bonis M., Connaître les émotions humaines, Mardaga, Hayen, 1996.

[18] Shrout P.E., Fleiss J.L., Intraclass correlations: Uses in assessing rater reliability, Psychol. Bull. 86 (1979) 420–428.

[19] Van Mechelen I., De Boeck P., Rosenberg S., The conjunctive model of hierarchical classes, Psychometrika 60 (1995) 505–521.

[20] De Boeck P., Rosenberg S., Hierarchical classes: Model and data analysis, Psychometrika 53 (1988) 361–381.