Eur Psychiatry 2002; 17: 1-10

S0924933802006843/FLA

ORIGINAL ARTICLE

Measurement of the perception of facially expressed emotions by a computerized device: method of analysis and research for the integration of emotions (MARIE)

Philippe Granato[1]*, Raymond Bruyer[2]

[1] *Centre Hospitalier de Valenciennes, Avenue Desandrouin, 59300 Valenciennes, France; Centre d'Investigation Clinique & INSERM, CHR&U, Lille, France;* [2] *Faculté de Psychologie et des Sciences de l'Éducation (unité de neurosciences cognitives), Université de Louvain-la-Neuve, Belgium*

(Received 30 August 2000; accepted 20 December 2001)

Summary – Current computerized tools allow detailed exploration of the structure and functioning of the "black box", i.e., human cognitive and affective systems as well as thought. This technology was used to study the visual perception of facially expressed emotions. Morphological transition from one canonical emotion to another led to the creation of a continuum of intermediary pictures, and the identification of perceived emotions by 65 normal subjects was measured. We call this application "MARIE" (in French: Méthode d'Analyse et de Recherche de l'Intégration des Émotions; Method of Study and Analysis of Integration of Emotions). Our study examined the relationship between the quantitative modification of the continuum and the quantitative variation of the responses. Standardization of graphs led to the assessment of the two parameters of a Laplace-Gauss curve, i.e., mean and standard deviation. It is argued that such a tool could be very useful in the clinical assessment of the emotional state of subjects and/or of patients. © 2002 Éditions scientifiques et médicales Elsevier SAS

Perception / Emotions / Facial expressions / Morphing / Laplace-Gauss

INTRODUCTION

Visual identification of facially expressed emotions is a frequent daily activity, and the ability to recognize faces and facial expressions is fundamental in social life [4,14,26]. The human brain seems to be innately able to discriminate facially-expressed emotions [19,35], and any failure of this ability often leads to difficulties in human relationships, especially for those people with some psychological frailness. However, identification of emotions is partly inferential, context-dependent and subjective, at least for non-basic emotions. Therefore, it is often uneasy to reach a consensus in a group of subjects who are asked to identify facially expressed emotions. People often express different, and sometimes opposing, viewpoints when observing the same stimulus. This observation is true in the case of perception of the light spectrum [3], auditory perception of phonemes [23,24], visual perception of facial identity [2] and of facial expression of emotions [6,7,18,20].

A direct study of this subjectivity is difficult and these processes are measured by means of (semi-) qualitative scales. Such scales cannot be accurately designed as long as this subjective dimension of detection by the viewer

*Corresponding author.
E-mail address: pgranato@nordnet.fr (P. Granato).

and/or by any control population applies. Nevertheless, this obstacle has been partly overcome by Ekman and Friesen [17]. Indeed, these authors identified “canonical” emotions, that is to say, a subset of emotional facial expressions for which a consensus was observed: joy, fear, disgust, anger, sadness, surprise and emotional indifference (neutrality). This consensus applies across and beyond age, race, nationality and cultural variations; when the control group is large, the consensus is even very close to 100%. However, unanimity is never reached. The difficulty of entirely cancelling out the subjective component from the perception of facial emotions probably led to the stagnation of progress in this field, which became, in turn, detrimental to the acknowledgment of its importance. Consequently, very few objective measurement tools are currently available, and still less tools are apt for use in a clinical setting.

Given the importance of emotions in both affective and relational social life and their implication in psychiatric and/or neurological conditions, the scientific study of emotions, including their subjective dimension, is highly relevant. The aim of the present study was to explore quantitative and measurable components of the perception of emotions, in order to propose a method of measurement enabling the comparisons between different groups of subjects and/or over the evolution of a particular illness.

Over the last decade, research in cognitive psychology [2,6,8,11,13,18,20,22,25,32,34], in cognitive neuropsychology [8,9], and in cognitive neuroscience [10,12] has used computerized morphing procedures to create continuous series of intermediate faces (morphs) between two original faces (sources). This design was used to study the perception of facial identity, facial gender and, more often, facial expression. Categorical perception of facial expressions was clearly demonstrated and is characterized by two features: (a) while morphs vary continuously between the two sources, the identification response varies in a sigmoidal manner (like a threshold function) revealing a boundary between two categories; (b) when pairs of morphs are displayed for same/different discrimination tasks, the response is easier and faster when the two stimuli derive from both sides of the boundary than when they issue from the same side. This robust observation solved the old controversy about the dimensional [29,30,33] vs categorical [15,16] nature of emotions.

The present study was aimed at examining the properties of the sigmoidal curve evidenced in the identification task, and to design a tool which would be easy to use in a clinical context. A group of young, healthy subjects were submitted to the procedure already described by the authors [6,20].

MATERIALS AND METHODS

Stimuli

Our previously published tool, MARIE (Méthode d’Analyse et de Recherche de l’Intégration des Émotions ; Method of Study and Analysis of Integration of Emotions), was used [6,20]. Method of study and analysis of integration of emotions is a computerized morphing program which computes the transition from one canonical expresssion A to another given canonical expression B (sources), by means of a continuum of 17 intermediate steps (morphs); these morphs and their sources are displayed at random for an identification response. Method of study and analysis of integration of emotions has been applied to the seven basic expressions (joy, sadness, fear, anger, surprise, disgust, neutrality) by using the photographs published by Ekman and Friesen [17], with permission (*figure 1*). All AB continua of sources are stored but, for the present study, only neutral/sad and angry/sad continua were selected (*figures 2a, b* display the series, and *figure 3*, a specific morph). Therefore, the material consisted of two series of 17 morphs each, plus the sources (size of the pictures = 10 × 18 cm; viewing distance = 40 cm). The selection of only these two series resulted from pragmatic considerations (duration), given the clinical goal of the study. In addition, it should be noted that the neutral/sad continuum allows the evaluation of the intensity of an emotional expression (sadness), while the angry/sad one captures the identification of the displayed emotion in a binary decision procedure. *Table I* shows the proportional contributions of A and B, respectively, in the computation of each morph.

In *table I*, it can be seen that the inter-morph step was not constant. Indeed, it was shorter around 50% than near the sources. The reason for this was to increase the sensitivity of the tool around the boundary area. Therefore, each morph (i) was characterized by a “transition value” (T_i) computed as $T_i = (2\ T_{i,b}/100) - 1$, where $T_{i,b}$ is the contribution of the source B to the morph i. In this way, a transition value varying from -1 to +1, with a null value for the central morph, was obtained. The right-most column of *table I* shows the transition values of the morphs.

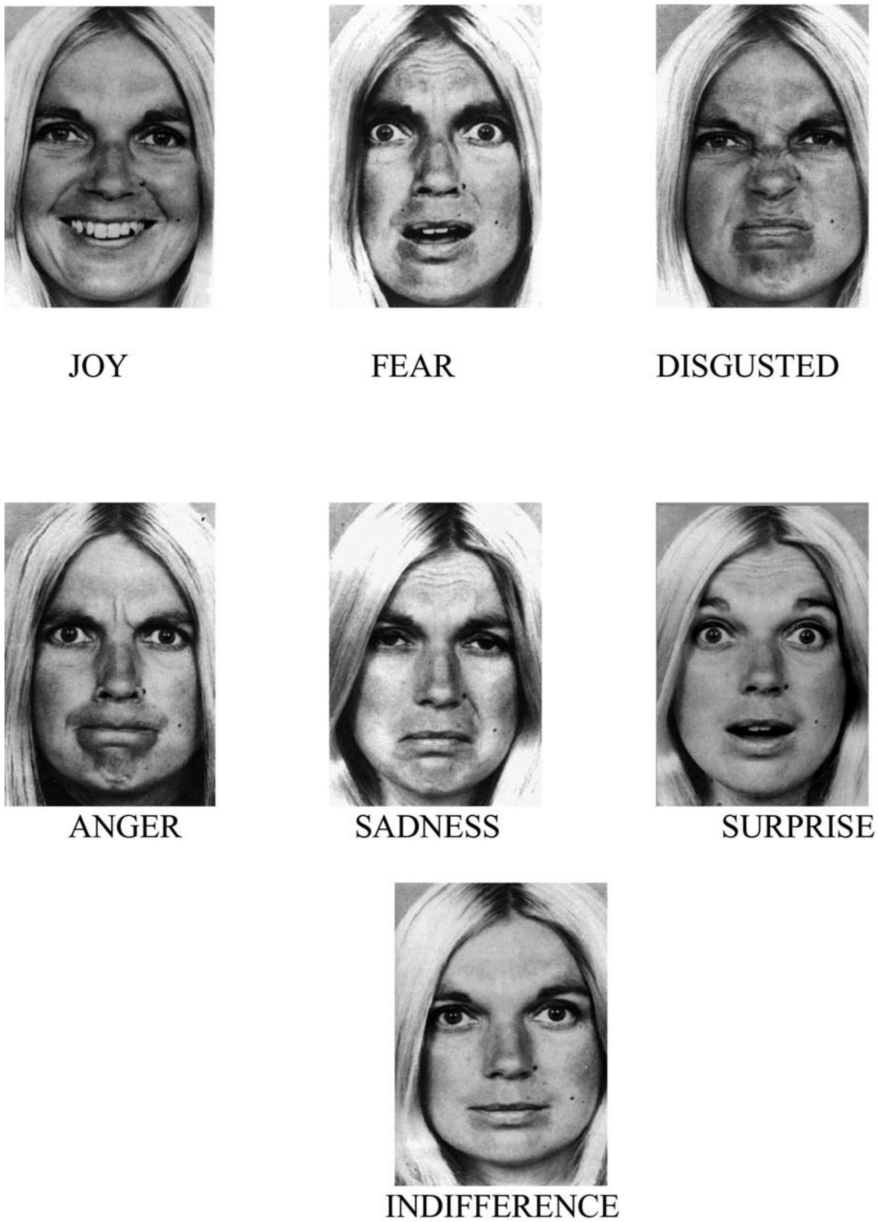


Figure 1. The seven emotional expressions used in the present tool (Ekman and Friesen 1975, with permission).

Subjects

Sixty-five right-handed adult subjects (43 females and 22 males), without any known current or previous neurological or psychological pathology, were recruited. In addition, a short psychiatric assessment was administered just before the beginning of the experimental session, so that only subjects in a normal mood state (according to the criteria of the DSM IV [1]) were enrolled. Visual acuity, whether corrected or not, was 10/10. Mean age was 32 ± 8 years (median 35.8; range 20–59). Educational level was high school ± 2 years.

Figure 2. The angry/sad (a) and neutral/sad (b) series used in the present study.

Method, design and procedure

All records were anonymous, each subject being identified by a serial number. The subject was tested individually, in a single session. The session began by the registration of information about age, gender, handedness, education, an informal assessment of medical and ophthalmological status, as well as a psychiatric evaluation. Then, the subject was asked to read aloud a short text on the screen of the computer used for the experiment in order to verify his/her normal visual acuity on this device. Finally, the experimental part of the session was introduced and familiarized by means of a task identical to the experimental one, but with stimuli resulting from the morphing of a circle with a square.

During the experimental part of the session, each morph was displayed on the screen of a portable computer (n = 38 stimuli). For each series, the 17 trials were randomly displayed (the same random order was used across subjects), and the series ended by the display of the two sources (trials 18 and 19). In each trial, the morph appeared at the center of the screen and was flanked by two verbal labels, defining the two canonical emotions of the series (*figure 3*). The subject had to identify the perceived emotion by choosing one of the labels, and to depress the left or right button of the mouse according to the left or right location of the label; this choice was mandatory and the morph remained on the screen until the response. Let us note that no correct (or erroneous) response was defined a priori.

RESULTS

Measurement of subject response

For each morph of each series, *table II* and *figure 4* display the number (and proportion) of subjects (out of

Figure 3. An example of stimulus (from the angry/sad series).

Table II. Number (and percentage) of subjects, out of 65, who chose the response "sad"

Pictures	Neutral/sad		Angry/sad	
	N	%	*n*	%
1	0	0.0	0	0.0
2	0	0.0	1	1.5
3	3	4.6	1	1.5
4	9	13.8	6	9.2
5	10	15.4	7	10.8
6	16	24.6	6	9.2
7	20	30.8	16	24.6
8	27	41.5	23	35.4
9	35	53.8	19	29.2
10	40	61.5	28	43.1
11	46	70.8	41	63.1
12	52	80.0	46	70.8
13	59	90.8	49	75.4
14	63	96.9	55	84.6
15	63	96.9	58	89.2
16	62	95.4	63	96.9
17	65	100.0	64	98.5
18	65	100.0	65	100.0
19	65	100.0	65	100.0

65) who chose the response "sad". Even though no error was defined a priori, some choices could be clearly discordant, i.e., different from those of the majority of the sample (outliers). So, for each morph *i*, the rate of such mistakes (M_i%) was computed as the proportion (in percent) of subjects who did not respond like the majority, according to the formula $M_i\% = 100 \min (n_{a,i}; n_{b,i})/N$, where $n_{a,i}$ and $n_{b,i}$ are the numbers of subjects who responded "A" and "B", respectively, and N the total number of subjects ($N = n_{a,I} + n_{b,i}$). In this way, M_i% was a measure of disparity and varied from 0 (unanimity) to 50 (high uncertainty, i.e., random choice at the sample level).

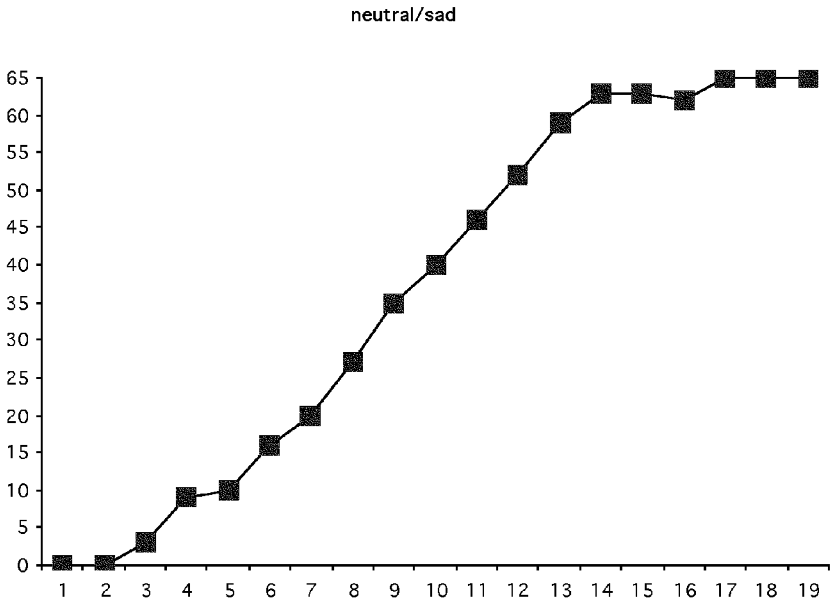
Table I. Description of the material

Pictures	Contribution of A (%)	Contribution of B (%)	Transition value (T_i)
1	100	0	–1.00
2	90	10	–0.80
3	80	20	–0.60
4	70	30	–0.40
5	65	35	–0.30
6	62	38	–0.24
7	59	41	–0.18
8	56	44	–0.12
9	53	47	–0.06
10	50	50	0.00
11	47	53	0.06
12	44	56	0.12
13	41	59	0.18
14	38	62	0.24
15	35	65	0.30
16	30	70	0.40
17	20	80	0.60
18	10	90	0.80
19	0	100	1.00

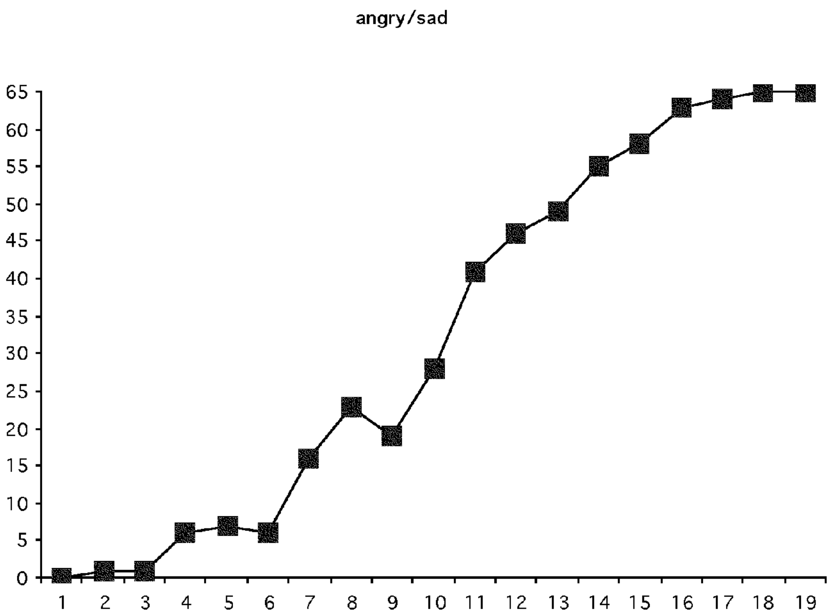
We focused on the function $M\% = f(T)$, that is to say, the disparity of choices according to the transition value of the morphs. This relation was analyzed as an observed probability density function and compared to a Laplace–Gauss distribution (whose parameters are the mean and the standard deviation). For each continuum, the deviation from the norm was computed by comparing the observed effects to the expected effects under the hypothesis of a normal distribution, by means of non-parametric χ^2 goodness-of-fit tests [31].

Results of analyses

For each continuum, the raw data (*table II*) were first compared to the results theoretically expected under the hypothesis of linearity, according to which the number of choices "B" would be linearly related to the



(a)



(b)

Figure 4. Graphical representation of the choices made by the subjects (a) neutral/sad; (b) angry/sad.

contribution of the source "B" in the morphs. Linear regressions were first computed for each series. For the neutral/sad series, the best fit of the observed data led to a straight line with a slope of 3.52 and an intercept of 0.36; in this condition, 67.3% of the subjects would perceive sadness in picture 10 (let us recall that the neutral/sad series actually displayed variations of intensity of sadness) and 50% of them would perceive sadness in a virtual picture with $T_i = -0.10$. For the neutral/sad series, the best fit of the observed data led to a straight line with a slope of 3.25 and an intercept of 0.05; in this condition, 47.4% of the subjects would perceive sadness in picture 10 (let us recall that, unlike the neutral/sad series, the angry/sad one actually displayed transitions between two different expressions) and 50% of them would perceive sadness in a virtual picture with $T_i = -0.02$.

However, for both series, the χ^2 test proved to be significant (neutral/sad: $\chi^2_{18} = 136.6$, $P < 0.0001$; angry/sad: $\chi^2_{18} = 136.6$, $P < 0.0001$), indicating that linearity was not an adequate fit of the collected data. Indeed, it can be seen in *figure 4* that the distribution of responses is much like a sigmoidal or threshold function. And accordingly, both tests were no longer significant when the data were submitted to the hyperbolic tangent transformation, i.e., a sigmoidal fit (neutral/sad: $\chi^2_{18} = 1.98$, NS; angry/sad: $\chi^2_{18} = 1.78$, NS).

Figures 5a (neutral/sad) and *5b* (angry/sad) display graphically the results transformed into M_i% values and expressed as a function of T_i. In the figures, the area under the part of the curve situated on the left (resp. right) side (where $T_i = 0$, i.e., the morph 50/50) indicates the number of subjects who identified the displayed emotion as “A” (resp. “B”).

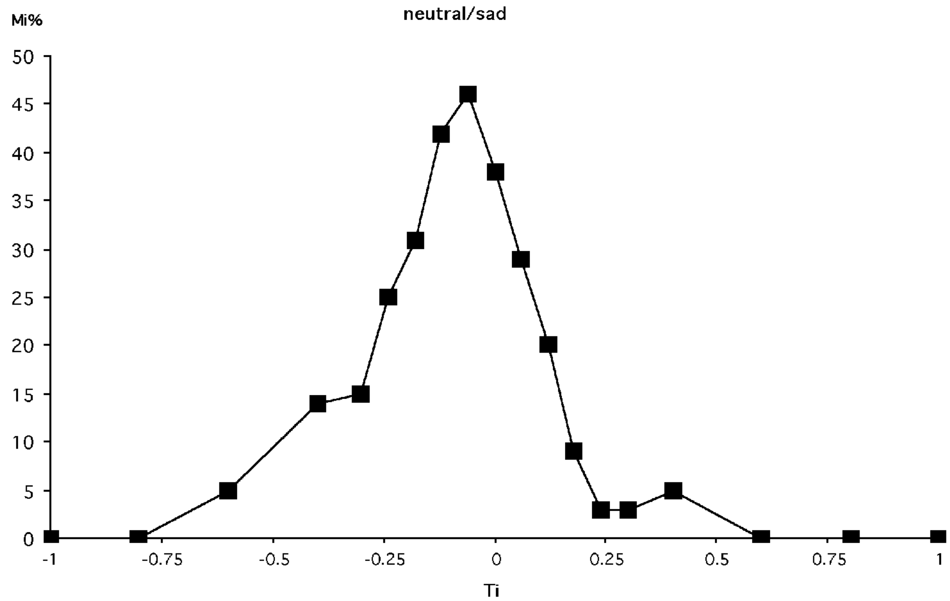
Descriptively, both curves look like the normal Laplace-Gauss distribution. And indeed, this visual impression was supported by statistical tests. A goodness-of-fit test of normality was computed for each graph and it was not significant in each series (neutral/sad: $\chi^2_{18} = 17.89$, NS; angry/sad: $\chi^2_{18} = 13.81$, NS). Thus, the series neutral/sad and angry/sad were well accounted for by a normal law. For angry/sad, the mean was – 0.032, very close to zero. This means that, around the 50/50 morph, the number of subjects who perceived anger was virtually identical to the number of subjects identifying sadness: subjects easily discriminated sadness from anger along the continuum. For neutral/sad, the mean was – 0.105. Thus, near the 50/50 morph, the number of subjects who perceived neutrality tended to be larger than the number of subjects who identified sadness.

The morphology of the two curves was similar, which is reinforced by the close values of their standard deviations (0.216 for neutral/sad and 0.241 for angry/sad). The standard deviation expresses the flatness of the curve, where a flat curve indicates a respectable rate of difficulty for the morphs which are relatively close to a source (or: far from the 50/50 morph). In this case, the mean is far from the sum (mean + standard deviation). Conversely, a peaked curve (small standard deviation) indicates no difficulty for morphs close to a source. In this case, the mean is close to the sum. It would be useful to know the normal distance between the mean and the sum, because the proximity of the observed standard deviations (0.216 and 0.241) could be, or not, due to chance and/or to specific characteristics of the recruited sample. This hypothesis requires the study of additional series of volunteers.

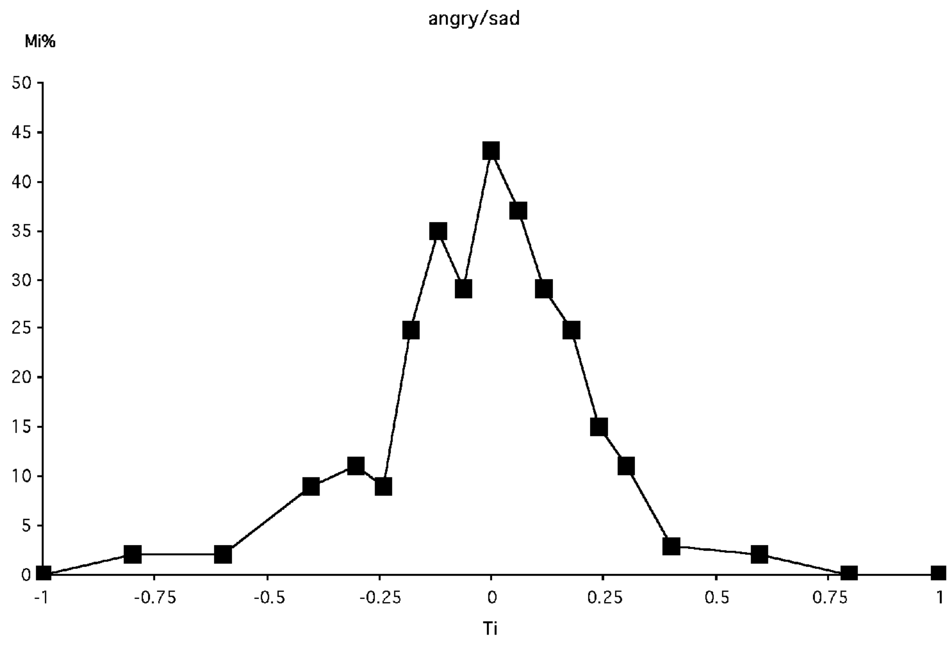
DISCUSSION

Subjects were shown two series of morphed emotional expressions for identification under a forced binary choice. The difficulty of choices progressively increased symmetrically as the central representation (morph 50/50) was approached. This observation is in good agreement with published reports dealing with emotional facial expressions [6,7,10,12,13,18,20,34]. Moreover, such difficulty in the visual interpretation of intermediary representations does not involve emotional expressions alone, as it has been evidenced for colors [3] and for visual identification of facial identities [2,22,25,32]. However, in these studies, it could be argued that decision making no longer had a binary aspect, as several responses were allowed at each trial. Therefore, the results of the present study could be an effect of the binary nature of choices. Two qualifications apply, however. First, the same pattern has been evidenced even when the identification of facial expressions is not binary [34]; second, this pattern has been observed for identification of face gender which is obviously a binary choice [11]. Consequently, the pattern of the present results was not a consequence of the binary nature of the allowed responses.

It is interesting to focus on this binary kind of choice, particularly on deviant responses (i.e., responses different from those of the majority of the subjects). Such “mistakes” can be measured individually—when a given subject doesn’t respond as the majority does—or collectively—as the proportion of subjects who do not respond as the majority does. Moreover, the association of a confidence interval (in which the effect is taken into account) would allow to compute a general measure of individual responses. Thus, the present design authorizes an experimental measurement of “mistakes” in the visual perception of emotions thanks to a binary, forced choice procedure. Moreover, it can lead to a numerical and graphical representation of “mistakes” as a function of the well-controlled continuum of morphs.



(a)



(b)

Figure 5. M_i% as a function of T_i, for the neutral/sad (a) and angry/sad series (b).

It would have been interesting to look for effects of gender on the distributions. However, as female subjects constituted two thirds of our sample (n = 22 males out of 65), such an analysis was not possible, and a complementary study is necessary to examine this specific question. However that may be, a tool as simple as the present one can be used in a clinical context and sufficed to replicate the published results [2,6,7,18,20].

Thus, the originality of this study lies in the simplicity of the procedure where the subjectivity of the subject, and therefore the "mistake", is the only thing taken into account. The definition of a normalized range –1/+1 on the *x*-axis (transition values of the pictures) and of a normalized range 0/50% on the *y*-axis (choices) leads to the design of a generalized graphical representation. In addition, the use of photographs for which a transcultural agreement is guaranteed [17] should make any comparison across studies easy, including the comparison of samples of patients suffering from the same or different conditions, without biases.

This standardized graphical representation was well fitted by a normal law. "Normal" is worth noting. Indeed, the identification of an emotional expression by a sample can be expressed by two parameters, mean and standard deviation, and the analysis of an emotion is not uniquely cognitive as control subjects take individual and personal decisions whose variations are reflected by the standard deviation. In other words, the interpretation of an emotion depends on both the stimulus and the mood state of the subject. In addition, the analysis of an emotion is simultaneously perceptual/cognitive and affective, that is to say, the participation in a test may induce a particular state of mood (for instance fear, anxiety or self-control), which may modulate the identification of emotional expressions. Nonetheless, this drawback can be turned into an advantage. Intuitively, it can be considered that a depressive subject will overestimate sadness and underestimate joy; and indeed, it has been shown empirically that the affective state of the perceiver modulates his/her perception of emotional stimuli [21,27,28]. So, the present tool would make it possible to evaluate the mood state of a subject by comparing his/her results to a baseline derived from the responses of a healthy, euthymic, sample.

As a final comment, it is worth considering facial expression of physical pain. Albeit not usually considered as an "emotion", this expression is clinically relevant. Tools like the present one could offer an objective and measurable quantification which may be useful to evaluate the analgesic effect of treatments, for instance.

MANQUE RENVOI DE BIBLIO

[5]

ACKNOWLEDGEMENTS

This work was supported by the grant 1998/1954 of the Programme Hospitalier de Recherche Clinique of the French governement. Thanks are due to Pascal Vincent, who orientated and guided this study with new and original insights and performed the statistical analyses; to Paul Ekman who gave permission to use photographs from "Unmasking the faces" (Ekman and Friesen 1975); to Olivier Lecherf who designed the computer program for processing and displaying the pictures; to the late Mrs Micheline Jankowski, director of the Institut de Formation en Soins Infirmiers du Centre Hospitalier de Valenciennes; and to Daniel Salvadori (Janssen laboratory). This study was made possible thanks to the Centre d'Investigation Clinique (prof. C.Libersa, CIC-CHU/INSERM, Lille).

REFERENCES

1 American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders. 4th ed. Washington, DC: ABA; 1994 (*French translation:* JD Guelfi et al. Mini DSM IV-critères diagnostiques. Paris: Masson; 1996).
2 Beale JM, Keil FC. Categorical effects in the perception of faces. Cognition 1995;57:217–39.
3 Bornstein MH, Korda NO. Discrimination and matching within and between hues measured by reaction times: some implications for categorical perception and levels of information processing. Psychol Res 1984;46:207–22.
4 Bruyer R. Le visage et l'expression faciale: approche neuropsychologique. Liège: Mardaga; 1983.
5 Bruyer R. Les mécanismes de reconnaissance des visages. Grenoble: Presses Universitaires de Grenoble; 1987.
6 Bruyer R, Granato P. Categorical effects in the perception of facial expressions: MARIE—a simple and discriminating clinical tool. Eur Rev Appl Psychol 1999;49(1):3–10.
7 Calder AJ, Young AW, Perrett DI, Etcoff ML, Rowland D. Categorical perception of morphed facial expressions. Visual Cognition 1996;3:81–117.
8 Calder AJ, Young AW, Rowland D, Perrett DI, Hodges JR, Etcoff NL. Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. Cognitive Neuropsychol 1996;13:699–745.
9 Calder AJ, Keane J, Cole J, Campbell R, Youg AW. Facial expression recognition by people with Möbius syndrome. Cognitive Neuropsychol 2000;17:73–87.
10 Campanella S, Hanoteau C, Dépy D, Rossion B, Bruyer R, Crommelinck M, Guérit JM. Right N170 modulation in a face discrimination task: an account for categorical perception of familiar faces. Psychophysiology 2000;37:796–806.
11 Campanella S, Chrysochoos A, Bruyer R. Categorical perception of facial gender information: behavioural evidence and the face-space metaphor. Visual Cognition 2001;8:237–62.
12 Campanella S, Quinet P, Bruyer R, Crommelinck M, Guérit JM. Categorical perception of happiness and fear facial expressions: an ERP study. J Cognitive Neurosci in press.
13 De Gelder B, Teunisse JP, Benson PJ. Categorical perception of

facial expressions: categories and their internal structure. Cognition and Emotion 1997;11:1–23.
14 Drag RM, Shaw ME. Factors influencing the communication of emotional intent by facial expressions. Psychonomic Sci 1967;8:137–8.
15 Ekman P. Universals and cultural differences in facial expressions of emotion. In: Cole J, editor. Nebraska Symposium on Motivation. Lincoln: University of Nebraska Press; 1972. p. 207–83.
16 Ekman P. Strong evidence for universals in facial expression: a reply to Rusell's mistaken critique. Psychol Bull 1994;115: 268–87.
17 Ekman P, Friesen WV. Unmasking the face. Englewood Cliffs (NJ): Prentice Hall; 1975.
18 Etcoff NL, Magee JJ. Categorical perception of facial expressions. Cognition 1992;44:222–40.
19 Field TM, Woodson R, Greenberg R, Cohen D. Discrimination and imitation of facial expressions by neonates. Science 1982;218:179–81.
20 Granato P, Bruyer R, Revillion JJ. Étude objective de la perception du sourire et de la tristesse par la méthode d'analyse de recherche de l'intégration des émotions « MARIE ». Annales Médico-Psychologiques 1996;154(1):1–9.
21 Halberstadt JB, Niedenthal PM. Emotional state and the use of stimulus dimensions in judgment. J Personal Social Psychol 1997;72:1017–33.
22 Levin DT, Beale JM. Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. Percep Psychophys 2000;62:386–401.
23 Liberman AM, Harris S, Hooman HS, Griffith BC. The discrimination of speech sounds within and across phoneme boundaries. J Exp Psychol 1957;54:358–68.
24 Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. Psychol Rev 1967; 74:431–61.
25 Livingston KR, Andrews JK, Harnad S. Categorical perception effects induced by category learning. J Exp Psychol: Learning, Memory and Cognition 1998;24:732–53.
26 Lockard J. Studies of human social signals: theory, methods and data. In: Lockard J, editor. Evolution of Human Social Behaviour. Amsterdam: Elsevier; 1980. p. 1–30.
27 Niedenthal PM, Cantor N. Affective responses as guides to category-based inferences. Motivation and Emotion 1986;10: 217–32.
28 Niedenthal PM, Halberstadt JB, Setterlund MB. Being happy and seeing 'happy': emotional state mediates visual word recognition. Cognition and Emotion 1997;11:403–32.
29 Plutchik R. The emotions: facts, theories, and a new model. New York: Random House; 1962.
30 Russell JA. A circumplex model of affect. J Personal Social Psychol 1980;39:1161–78.
31 Schwartz D. Méthodes statistiques à l'usage des médecins et des biologistes. Paris: Flammarion; 1975.
32 Stevenage VS. Which twin are you? A demonstration of induced categorical perception of identical twin faces. Br J Psychol 1998;89:39–57.
33 Woodworth RS, Schlosberg H. Experimental psychology, revised edition. New York: Holt; 1954.
34 Young AW, Rowland D, Calder AJ, Etcoff NL, Seth A, Perrett DI. Facial expression megamix: tests of dimensional and category accounts of emotion recognition. Cognition 1997;63: 271–313.
35 Young-Browne G, Rosenfeld HM, Horowitz FD. Infant discrimination of facial expressions. Child Dev 1977;48:555–62.