DYNEMO: A VIDEO DATABASE OF NATURAL FACIAL EXPRESSIONS OF EMOTIONS

Anna Tcherkassof[1], Damien Dupré[1], Brigitte Meillon[2], Nadine Mandran[2], Michel Dubois[1] and Jean-Michel Adam[2]

[1]LIP, Univ. Grenoble Alpes, BP 47 - 38040 Grenoble Cedex 9, France
[2]LIG, Univ. Grenoble Alpes, BP 53 - 38041 Grenoble Cedex 9, France

ABSTRACT

DynEmo is a database available to the scientific community (https://dynemo.upmf-grenoble.fr/). It contains dynamic and natural emotional facial expressions (EFEs) displaying subjective affective states rated by both the expresser and observers. Methodological and contextual information is provided for each expression. This multimodal corpus meets psychological, ethical, and technical criteria. It is quite large, containing two sets of 233 and 125 recordings of EFE of ordinary Caucasian people (ages 25 to 65, 182 females and 176 males) filmed in natural but standardized conditions. In the Set 1, EFE recordings are associated with the affective state of the expresser (self-reported after the emotion inducing task, using dimensional, action readiness, and emotional labels items). In the Set 2, EFE recordings are both associated with the affective state of the expresser and with the time line (continuous annotations) of observers' ratings of the emotions displayed throughout the recording. The time line allows any researcher interested in analysing non-verbal human behavior to segment the expressions into small emotion excerpts.

KEYWORDS

Facial Expressions, Dynamic Database, Spontaneous Behavior, Emotion Communication, Emotion Perception, Continuous Annotation

1. INTRODUCTION

Emotion is an essential component of interpersonal relationships. Hence, the study of the affect underlying human behavior is central to research concerning interaction processes. This problematic is especially relevant in the field of social sciences [1]. Some researchers interested in the analysis of human behavior during interactions are more particularly interested in the question of emotional facial expressions (EFEs). For instance, during the past few years, considerable effort has been devoted to developing tools for EFEs analysis. Indeed the potential utility of a system capable of analyzing facial expressions automatically and in real time is considerable in terms of its possible applications (remote detection of people in trouble, detection of malicious behavior, multimedia facial queries, etc. [2]. Automatic analysis of EFEs requires facial display databases for use both as learning resources and for testing automatic recognition systems [3, 4]. In response to the needs of the scientific community of behavior analysis such as emotion databases for example (for a chronological review, see [5]) . The issue to record and to analyze ecological behaviors is especially important for affective studies concerned with realism, notably for pragmatic matters [6-8]. It is the case, for example, of the field of automated systems. To be reliable, such systems must be trained on facial patterns close as possible to naturally

occurring human affective behavior because facial expression analysis must be able to distinguish with accuracy the everyday emotions expressed by people. Undeniably, there are fundamental differences between non-natural behavioral stimuli and those exhibited in everyday life. Research on facial expression has highlighted how dynamics and spontaneity are crucial expressive features of natural displays.

Firstly, in a natural approach to human interactions, the dynamic quality of EFEs is also an essential feature. For instance, researchers found that deaf accuracy for facial expressions sign language depends upon the use of material that preserves the dynamic changes occurring during the unfolding of a facial expression [9]. In addition, dynamic displays facilitate facial expression recognition by autistic children, contrary to what is usually found when they are presented with static faces [10]. Indeed, EFEs displayed in real life are dynamic and change rapidly from one expression to another. Thus, the motion of facial expression provides much more information than that provided by static expressions: time course [11], change of speed [12-14], facial-feature amplitude [15] and rate of change, expression's irregularity and intensity [16] to name a few. Still expressions do not present such information and leads to facial recognition differential effects [17-19]. Morphed stimuli are not satisfactory either due to the question of real world validity [20]. As a matter of fact, morphing methodology conceives facial-expression changes in a linear fashion, whereas natural expressions have unique trajectories resulting from a combination of different patterns each with its own dynamic. For instance, dynamic face material contains an apex for each facial feature in motion [21]. Dynamics may be irrelevant for work in automated facial expression recognition seeking frame-by-frame classification. However the tendency is now to automatic motion analysis [22]. For this kind of research, video databases are necessary. Yet, such databases should also supply natural recordings, which is seldom the case. For instance, the GEMEP or the ADFES databases provides videos of facial recordings (i.e. dynamic EFEs), but the displays are not spontaneous.

Secondly, spontaneously occurring behavior differs in various aspects from deliberate behavior, including timing, visual appearance [23, 24]. In particular, some spontaneous facial configurations can hardly be performed deliberately [25]. For instance, even if it seems rather easy for anyone to raise the outer corner of eyebrows spontaneously while keeping the inner brows at rest, few can perform this facial action voluntarily. Another example is the spontaneous facial expression of sadness usually characterized by a depression of the lip corners and a rising and narrowing of the inner corners of the brows. Yet, few people can perform these actions deliberately without training. Moreover, spontaneous facial actions have different temporal and morphological characteristics (duration, intensity and asymmetry) than posed ones [26, 27]. Many spontaneous smiles are longer in total duration, slower in onset and offset, smaller in amplitude, and are more symmetrical than posed smiles [28-30]. As realistic as expressive stimuli generated by actors may be, they cannot provide the same information as natural affective expressive behavior [31]. From an ecological perspective, affective analyses based on deliberate EFEs have a poor generalization capacity, which is why they need examples of naturally expressed EFEs and not prototypical patterns of facial behaviors. They need methodologies providing recordings that are truly spontaneous and authentic [3]. However such databases are still rare because of the practical and ethical difficulties of recording natural EFEs.

Lastly, the reliability of affective studies depends on another crucial issue, namely, the content analysis. Most affective behaviors that occur in realistic interpersonal interactions or in human-computer interactions are not basic emotions but pervasive ones ("forms of feeling, expression and action that color most human life", *sic* [32]) . Recent studies on natural facial expressions have even begun to look at mental states [33] such as interest, boredom, anxiety, and thoughtfulness, which are the kinds of expressions most often observed in face-to-face and man-machine interactions. Many efforts have been done toward the collection of such emotional behaviors. But most existing natural databases have the inherent flaw of not specifying the

emotional state represented by each EFE recording [3]. In other words, methods to measure and to analyze affective behaviors remain to be provided.

The term affective behavior refers to two distinct entities. The first is the internal state that produced the behavior, i.e., the emotion experienced by the person whose EFEs are being recorded (called the encoder). The second concerns the observer's perception of the emotion (or emotion interpretation, recognition, or attribution). The distinction between production and perception of faces must be clearly defined because the two entities relate to different research questions [34]. Ideally, an affective behavior database would supply natural and dynamic stimuli accompanied by an objective specification of the encoder's emotional experience. The recordings should also be accompanied by the meanings observers find in the encoder's behaviors. Last but not least, [3] deplore the lack of contextual variables describing the recordings. In relation to the production/perception issue, metadata about the emotion-inducing context (type of environment, what the encoder is doing when filmed, possible presence of others, etc.) must be specified in descriptions of emotional content. Such metadata are especially necessary due to the strong context dependence of emotions [35, 36]. So databases must accurately specify the emotional content of the recordings they contain, both in terms of production, perception and contextual information.

In the present article, we present DynEmo which compiles the key features needed for a high-quality database. It aims to provide encoding and decoding method to measure and analyze affective behaviors. Its purpose is to add to the range of naturalistic affect resources [37-39]. Here, encoders are ordinary people and the age range of the encoders is wide (from 25 to 65 years old). DynEmo's objective is to provide spontaneous and dynamic affective behaviors supplemented with data regarding the emotional content of the recordings (production *versus* perception levels), as well as contextual information. It was developed by a multidisciplinary team of psychological researchers, computer scientists, statisticians, experimentation and instrumentation specialists, and a legal professional, who worked together closely throughout the various stages of the project. The final corpus intends to meet all psychological, technical, and ethical criteria. The procedure used to develop DynEmo is described below.

2. ENCODING SPONTANEOUS FACIAL EXPRESSIONS (SET 1 AND 2)

The database was elaborated in two Sets: the Set 1 with only the encoding phase and the Set 2 with both the encoding and the decoding phase. For the encoding phase (Set 1 and 2), since pervasive EFEs are hard to collect, the main challenge was to record affective behaviors in realistic but controlled emotion-eliciting conditions. The authenticity of the recordings is ensured by the elicitation method, which guarantees to some extent that the elicited affective displays are spontaneous and not posed. Gathering such natural facial displays implies selecting standardized tasks designed to elicit affective states, and having encoders filmed without knowing so to make sure they behave authentically. This also implies that encoders are not aware that they are carrying out an emotion-eliciting task. For this purpose, each encoder was recruited by a private company for a study supposedly devoted to an "ergonomic visual task". Via this "cover story", the encoder was led to believe that the focus of the research was on something other than the facial expression of emotion. In addition, each encoder performed only one elicitation task to conceal the real purpose of the experiment. Performing only one task has the additional advantage of preventing any emotional "pollution" (i.e., contamination of an emotion elicited by one task upon the elicitation of another affective state by a different task). The encoding phase also dealt with the assessment of the affective state of the encoder generated by the emotion elicitation, in order to characterize the affective content of the recorded EFEs.

2.1 Data acquisition

The recording conditions (camera parameters, camera distance, and illumination setting) were carefully controlled. The video recording procedure went as follows. In the recording room the screen of a PC was projected on the wall using a video projector (Figure 1). The encoder sat on a chair at a small table facing the wall, and could use the keyboard and the mouse of the PC when required by the elicitation tasks. A set of speakers broadcasted the sound of the multimedia material displayed to the encoder, and another speaker was used by the researcher in the control room to broadcast some comments needed during certain tasks. Two cameras were hidden in the room to keep the encoder from knowing s-he was being filmed. A Pan Tilt Zoom camera located under the projection screen filmed the whole face of the encoder. Another camera, positioned 45 degrees to the left of the encoder behind a folder on a shelf, filmed the encoder in full. A clip-on microphone was fastened under the table so that it was possible to listen to the encoder from the control room. During the encoding session, the encoder was left alone in the room (except for Task 10, see Table 1). Once instructions for the concerned computer task were given to her-him by the researcher, the encoder was left alone in the room. When finished, s-he was given a 15€ coupon for her-his participation after being debriefed.
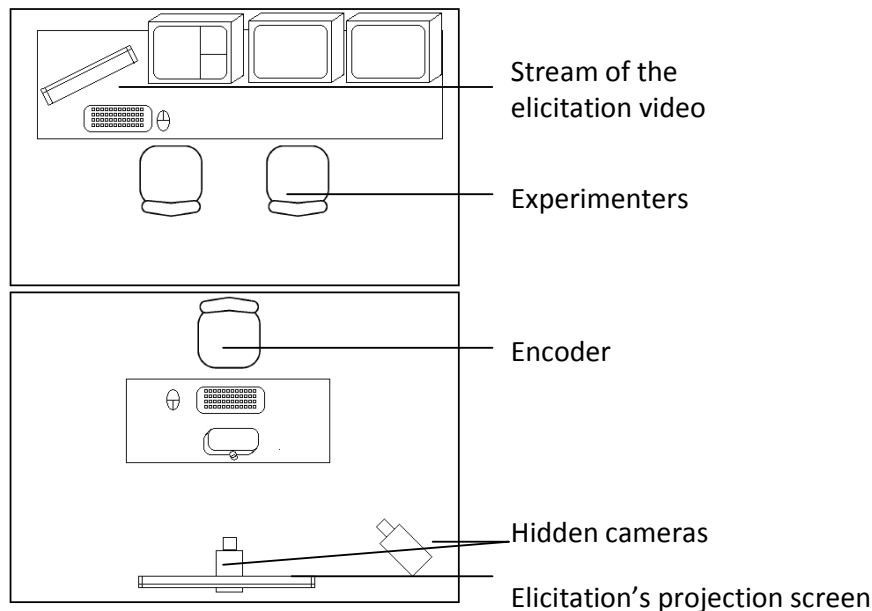
Figure 1. Design of the technical and recording rooms

The context in which the encoders' EFEs were produced is well documented. Metadata about the filming context is available, in particular, what the encoder was doing and the type of information s-he was reacting to (multimedia material described below). The metadata are supplied by mixing three streams (the stream of the encoder's screen displaying the emotion elicitation video, the zoomed-face stream, and the task stream, see Figure 2).

Figure 2. Three video streams gathered and mixed in real time
(encoder's zoomed in face, full length and screen activity)

In all, the database corpus contains the EFE records of 358 ordinary Caucasian participants (182 female encoders and 176 male encoders, ages 25 to 65 years old, M = 48, SD = 9.21, see Table 1).

Table 1. Encoders' gender and age distribution according to eliciting tasks

Elicitation	Female					Male					Total
Age	20-29	30-39	40-49	50-59	60-69	20-29	30-39	40-49	50-59	60-69	
Positive task 1	0	4	7	10	0	0	4	6	7	1	39
Positive task 2	0	5	7	7	2	0	4	7	7	1	40
Positive task 3	0	4	7	10	1	0	4	6	6	3	41
Positive task 4	0	3	7	5	5	0	4	6	6	2	38
Negative task 5	0	5	7	9	3	0	3	5	3	5	40
Negative task 6	2	2	6	7	4	0	6	6	5	4	42
Negative task 7	0	3	9	6	3	0	4	6	4	7	42
Negative task 8	0	4	7	5	2	0	4	7	4	2	35
Negative task 9	0	0	3	1	0	0	5	7	3	1	20
Negative task 10	0	2	1	3	0	0	2	0	1	1	10
Neutral task	0	1	2	1	0	1	1	2	3	0	11
Total	2	33	63	64	20	1	41	58	49	27	358

2.2 Ethical Issues

This protocol raised several ethical problems. The first one concerned the hidden cameras and the use of a cover story, necessary for scientific and ecological purposes, but needing to match deontological and juridical requirements. Another issue was the question of respect towards the encoder's privacy and image reproduction rights. Indeed, obtaining the consent of encoders is required to make data publicly available. These problems were handled by a jurist involved at the beginning of the project before starting the encoding sessions. To overcome these ethical and juridical obstacles encoders signed two consent forms. The first form was given to participants before the encoding session began and asked for their agreement in using the data issued from the test for research purposes. Only very general information was provided to keep hidden the real purpose of the study (information such as upholding of the participant's data privacy, no commercial use of the data, contact address for further information, etc.). It was notified that they could stop the experiment at any moment. The second form provided to participants after the

debriefing session was far more detailed. This form was specifically concerned with the use of video during the session to be used in building a database for researchers. Encoders were informed that they could rescind their consent at anytime. Few of them refused to sign and, when they did, all their materials were instantly destroyed. The agreements of encoders thus make the present data publicly available.

Another challenge was the necessity of keeping anonymous the questionnaire data, considered private information from a juridical point of view, in conjunction with the possibility to subsequently remove and destroy all materials — in case an encoder rescinds her/his consent someday. A specific coding nomenclature for all the files coming from the encoding sessions, excluding any name, was adopted, so that it was impossible to associate *a posteriori* an encoder to her or his personal data from any computer with DynEmo data. This procedure allows the subsequent removal and destruction of the data from the corpus if necessary while guaranteeing anonymity.

Finally, during the debriefing session, it was verified that encoders were not durably affected by the emotional induction. The workings and aims of the database were also carefully explained to participants (all full recordings of the debriefing contents are available on request). Psychologists researchers made certain the participants were in a state of serenity upon leaving the session.

2.3 Elicitation Tasks

In order to record pervasive facial behaviors in a highly controlled environment, the computer tasks were chosen as the emotion inducing circumstances. Four positive elicitation tasks, six negative, and one neutral elicitation tasks were developed (Table 2) to satisfy the need to quickly elicit different pervasive emotional states that varied in intensity (except for the neutral task). They were also designed in such way that the encoding context did not explicitly evoke the question of emotion. In this way, the encoder was not led to put on deliberate facial expressions. The EFE elicitation tasks consisted of slide shows of videos, pictures and text, and did not require any action from the encoder (passive tasks). There were also some active tasks, some of which used materials based on the "Wizard of Oz" technique (e.g., Task 9). The materials and further descriptions are available upon request.

Once the emotion-eliciting task was over, the encoder immediately filled out three questionnaires regarding her-his emotional state during the task. The encoder's affective state was assessed using three emotion measures: dimensional items [40], action-readiness items [41], and emotional labels items. Dimensional items and action-readiness items (which are less explicit affective measures than emotional labels) were assessed first in order to avoid overly leading questions about the encoder's emotional state from the beginning of the assessment session. Thus the first questionnaire concerned the measure of the affective state through dimensional scales [40] (e.g. *The state you feel after carrying out this visual task is: Sleepiness* [to] *Aroused*). The second questionnaire measured affective state through action readiness scales [41] (e.g. *The visual task you just carried out stirs up a tendency to approach, to make contact*), and the third one measured the emotional state using 12 emotional labels scales (e.g. *How much this visual task made you feel frightened?*). For this measure, pervasive labels were preferred to basic ones to avoid any theoretical confusion. For each item the encoder answered using a 6-points scale. An ending open item allowed the encoder to write down any other emotion label s-he considered more appropriate. The 3 questionnaires provide a thoroughly indicator of the affective state of the encoder engendered by the computer task (Appendix 1). Answers to the questionnaires were recorded on-line and were linked to the encoder's recording via a specific nomenclature [42]. At the end of the questionnaire, the researcher came into the room for the debriefing session.

Table 2. Affective-inducing computer tasks.

Emotion Task	Length	Task Description
Positive task	*	The encoder sorts 4 funny TV advertisements according to her/his preferences.
Positive task 2	00:04:07	The encoder views a slideshow of optical illusions and their explanation, followed by a TV commercial for Audi cars illustrating these optical illusions.
Positive task 3	*	The encoder is asked to evaluate, as accurately as possible, the volume of each of 20 containers displayed without any helpful cues regarding their size. Whatever her-his performance, s-he is ultimately congratulated for having successfully performed the task as compared (supposedly) to other encoders.
Positive task 4	*	The encoder is shown an attention-biases video (Simons & Chabris, 1999) for which experimental psychology has shown that 75% of viewers discover (usually with surprise) that they had not noticed a gorilla crossing the scene they had just seen while focusing on a counting task.
Negative task 5	00:01:04	The encoder watches a video report of a surgical operation of a firefly larva located under a woman's scalp.
Negative task 6	00:01:45	The encoder is displayed a slideshow about human misery, based on the contrast between healthy, smiling children and "soldier" children, followed by land-mine-wounded children, very poor children, and starving children. Opera music accompanies the slideshow from the beginning until the pictures of the poor children are shown. The rest of the slideshow had no music, making it morally more burdensome.
Negative task 7	00:02:28	On a screen with a white background, 8x6 black spots turned red one after another, from top left to bottom right. The time between one color change and the next is lengthened by 4 ms each time (2s between the first two changes 20 s between the last two spots).
Negative task 8	00:00:20	The encoder is shown a German TV commercial for an energetic drink in which a car is moving at a steady pace along a winding road. The camera lens is following the car and the encoder hears soft low-volume music. The car disappears behind a few trees, but the camera continues to move. At the moment when the car is expected to reappear, a zombie face suddenly appears with a shrill scream. The music is replaced by a loud heartbeat.
Negative task 9	*	The encoder is asked to evaluate, as accurately as possible, the volume of each of 20 containers displayed without any helpful cues regarding their size. S-he gives her-his answer in a loud voice to a (fake) defective automatic speech-recognition system. The poor performance of the system compels the encoder to repeat her-his answers again and again.
Negative task 10	*	The encoder is set with a heartbeat monitor on her-his chest and the researcher (a woman for male encoders; a man for female encoders) remains in the room during the entire encoding session. The encoder is led to believe that s-he is physiologically aroused by pictures of naked and/or provocative by posed persons (women for male encoders; men for female encoders), as suggested by the heartbeat broadcasted from the loudspeaker. The fake heartbeat accelerates when the picture is assumed to be exciting and decreases when it is not (the pictures of persons in everyday situations like riding a bicycle, talking to another person in the office, etc.).
Neutral task	00:00:06	A six seconds animated (spots changing color) screen.

2.4 Effectiveness of Elicitation Tasks

2.4.1 Dimensional and action-readiness scales
The 10 elicitation tasks succeeded in inducing an affective state as compared to the neutral task (Table 2, see also all the results are presented in Appendix 2). For instance, negative tasks 5, 6, and 8 were rated as inducing a significant negative valence (M_5 = 3.15, M_6 = 2.55, and M_8 = 2.63,

respectively). Negative tasks 6 and 8 were also rated as significatively inducing discomfort (M_6 = 2.93 and M_8 = 2.91, respectively). Positive tasks 3 and 4 were rated as significatively inducing no A*gonistic* and *Inhibiting* action tendencies (A*gonistic*: M_3 = 1.24 and M_4 = 1.42; *Inhibiting*: M_3 = 1.41 and M_4 = 1.62). Negative tasks 6 and 8 were rated as significatively inducing *Not to exist* (M_6 = 5.17), *Not to laugh* (M_6 = 1.26), *Crying* (M_6 = 2.9) action tendencies and *Protection* (M_8 = 3.2), *Rejection* (M_8 = 4.51), and *Tension* (M_8 = 2.23) action tendencies.

2.4.2 Emotional label scales

Overall, 85.2% of the encoders indicated having moderately or strongly felt an emotion among the labels proposed on the third questionnaire (Table 3). Out of 358 encoders, 325 indicated that they felt at least one affective state strongly (rating of 5 or 6 on one emotion-label scale, M = 5.46, SD = 0.78). The tasks varied as to the kind of emotion elicited. For example, 93% of the encoders on task 6 reported having experienced an emotion. Most said they were moved by the task (69%). High percentages were also observed for task 2: 90% of the encoders reported having experienced an emotion (with an overwhelming majority of 90.5% describing their emotional experience as interested/curious). Regarding task 8, 94.3% of the encoders reported having been elicited emotionally: 70.9% reported experiencing fright, the other elicited emotion was astonishment (63.3%). Overall, for all tasks pooled, amusement, astonishment, and curiosity were the affective states most often reported as being strongly experienced, while disappointment and boredom were reported the least often.

Table 3. Percentage of encoders who indicated having experienced moderately or strongly (5 or 6 respectively, on a 6-point scale) at least one affective state on the emotional label scales

Task	at least 1 affective state	Curiosity	Amusement	Astonishment	Satisfaction	Boredom	Disgust	Fright	Moved	Shame	Annoyance	Disappointment	Humiliation
Positive task 1	90%	85%	62%	15%	8%	3%	5%	3%	10%	0%	0%	0%	0%
Positive task 2	90%	90%	18%	28%	13%	0%	0%	0%	10%	0%	3%	0%	3%
Positive task 3	90%	85%	32%	12%	15%	0%	0%	0%	5%	0%	0%	0%	0%
Positive task 4	90%	84%	29%	16%	8%	3%	0%	0%	0%	0%	0%	0%	5%
Negative task 5	98%	83%	3%	45%	0%	3%	35%	20%	25%	0%	3%	0%	3%
Negative task 6	93%	41%	2%	14%	0%	0%	17%	26%	69%	38%	17%	0%	12%
Negative task 7	86%	83%	7%	29%	12%	14%	0%	0%	5%	0%	0%	2%	2%
Negative task 8	94%	69%	9%	63%	6%	3%	26%	71%	26%	3%	17%	3%	14%
Negative task 9	80%	70%	35%	10%	10%	0%	0%	5%	5%	0%	20%	10%	0%
Negative task 10	80%	60%	10%	0%	0%	0%	0%	0%	0%	10%	0%	0%	10%

3. DECODING SPONTANEOUS FACIAL EXPRESSIONS (SET 2)

The decoding phase dealt with collecting measures of the emotional content displayed in the EFE recordings. More specifically, these measures consisted of having observers make judgments assessing the emotional content of each recording with by making dynamic ratings.

Therefore, this section focuses on analyzing the emotion annotation method. 125 EFEs recordings (Set 2) were continuously rated by observers (hereafter called decoders) using a computer-based tracing tool called *Oudjat* [23]. *Oudjat* is a data-processing interface that provides dynamic ratings. It allows decoders to assess online the emotions they perceive in an encoder's face. It collects participants' judgments as they view videos, recording their judgments every 0.1 sec.).

Responses are given in two stages (Figure 3). The first stage consists, for the decoder, of delineating the emotionally expressive parts of the EFE recording. The second stage consists of indicating, for each emotionally expressive part previously delineated, the emotion the encoder is judged to be feeling.
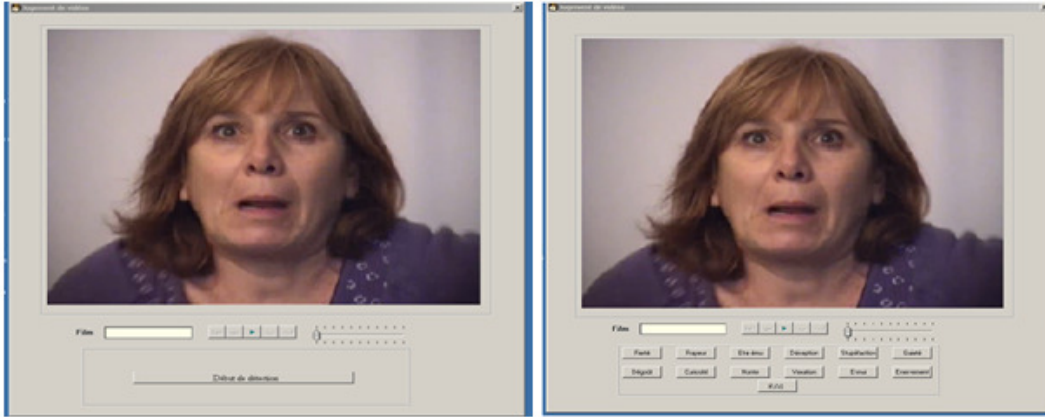

Figure 3. Screenshots taken from the *Oudjat* annotation procedure.

Here the *Oudjat* assessment task procedure was as follows. On the first step, a series of EFE recordings was randomly displayed to the decoder on a computer. As each EFE recording was being played, the decoder "marked" the video by pressing a key each time s-he perceived the beginning or the end of an emotion displayed on the encoder's face. For each recording, several expressive segments could be identified. An algorithm prevented the first step from exceeding ten minutes. It also prevented an EFE recording to be displayed twice to a decoder. Other than that, the EFE recordings were displayed randomly. Thus, each decoder was shown one to ten videos (according to their length). On the second step, the software replayed the previously marked segments of each EFE recording and stopped after each segment. The decoder then assessed this expressive segment by selecting one of the 12 proposed emotional labels in a drop-down menu. The emotional labels were the same as those used by the encoders for their emotional self-reporting. The whole assessment session lasted about 30 min.

426 decoders have rated Set 2 using *Oudjat*. The decoders are undergraduate students at the University of Grenoble, France, with no experience in facial expression annotation. Each of the 125 EFE recordings has been assessed by about 20 decoders. The number of decoders per video varies due to the randomized selection of the videos.

3.1 Emotional Ratings of the EFE Recordings

For each EFE recording, *Oudjat* provided a timeline showing the emotional label selected by each decoder for each marked segment. A general timeline was then calculated by superimposing, every tenth of a second along the entire EFE recording, the timeline of all decoders. This overall timeline shows the evolution (superimposed curves) of between-decoder agreement. In other words, it shows all participants' judgments synchronized with the real-time unfolding of the EFE. For example Figure 4 consists of an EFEs recording of a female encoder during task 5. During the first step of annotation, 100% of decoders have segmented the recording from sec. 34 to sec. 54 as being emotionally expressive. During the second step, 70% identified it as expressing disgust (in blue) whereas 30% have rated the face as expressing fright instead (in yellow). Thus the decoded emotions were identifiable readily.
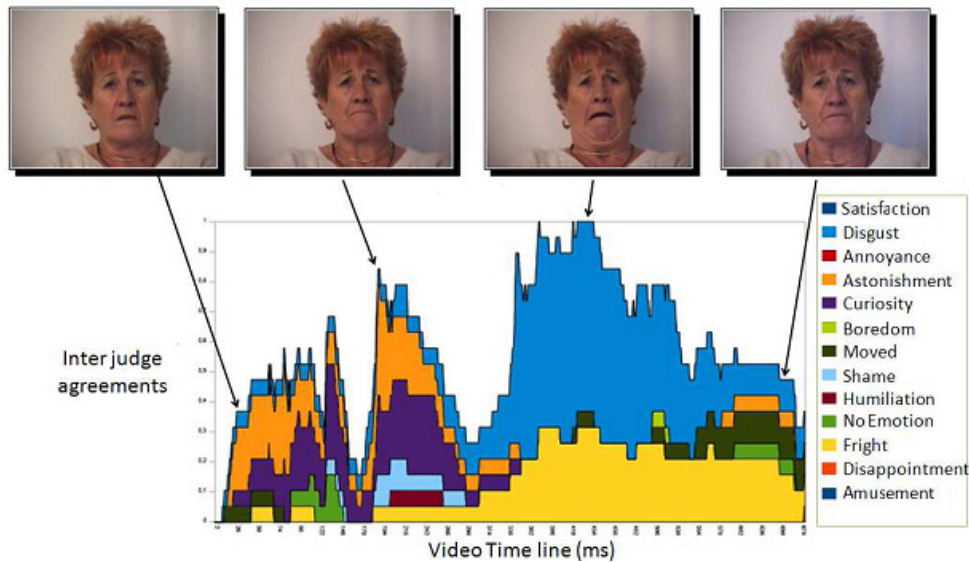
Figure 4. Emotional expressive time-line. Frames are taken from the video of an encoder who reported disgust and its corresponding underneath time-line

Because these ratings are dynamic, they cannot be compared with those obtained from other emotional datasets. Instead, we computed emotion identification scores. This was done for each EFE recording by determining which frame garnered the most inter-judge agreement on a given emotion label. Emotion identification scores are portrayed in Appendix 3. These scores indicate, for each segment, the emotional label selected the most by decoders. As is often the case, some facial expressions achieved an extremely high decoding agreement rate whereas others were poorly identified [43, 44]. Nevertheless, on average, there was a substantial amount of agreement between decoders. All measures are available online as support material.

3.2 Analysis of Inter-Decoder Reliability

The reliability of the marked segments was estimated using intraclass correlations (ICC, cf. Appendix 3). The calculation formula for the ICC is adapted from the notation used in Bland and Altman [45]:

$$ICC = [MS(B) - MS(W)] / [MS(B) + (m-1) MS(W)]$$

where MS(B) and MS(W) are the mean squares for between subjects and within subjects, respectively, and m is number of observations per subject [46].

On average, the ICC for the 125 videos is .79 (*SD* = .18), which is a good reliability level. In other words, EFE recording segmentation was, as a whole, homogenous across decoders. They agreed as to when a face was emotionally expressive and when it was not. All ICC values are available online as support material.

4. DYNEMO END USERS

A website hosts the DynEmo database. The EFE recordings are intended for scientific research purposes only. They can be downloaded or displayed in web streaming for free at the following address: https://dynemo.upmf-grenoble.fr/. To obtain a login on the site and download the data, a consent paper must be signed. An English version is underway. The website allows the user to

simultaneously watch both an EFE recording and its timeline (Figure 5). Self-reported emotions data are supplied in a separate document are also available.

The database contains 358 EFE zoom video files (for the entire duration of the elicitation task) with a 768 x 576 pixel resolution (25 fps). The file format provided for videos is MPEG4. Video files were encoded using H264 Codec.

EFEs recording selection and encoder filtering occur online. Several drop-down menus listing different variables (type of emotion, sex of encoder, age, etc.) allow the user to specify the kind of recordings s-he is interested in. A table displays all of the EFE recordings that match the request, along with their affective description (i.e., emotions felt by the encoder, see Index Card I in the Appendix 1, and general timelines, see Figure 5).
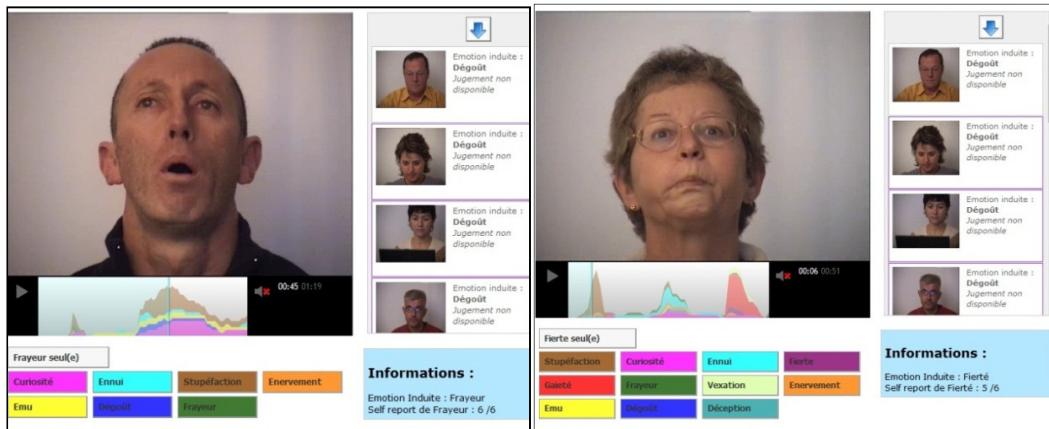

Figure 5. Screenshots taken from the DynEmo database.
The timeline is synchronized with the progression of the recording.

Users download the whole recording. Though questionable for persons directly interested in its most expressive parts, this procedure has the advantage of allowing someone else interested in the non-prototypical displaying of emotions to gain access to this kind of material. Also, because of the impact of idiosyncratic features on emotion recognition, a researcher can look for an expressive segment preceded by a (more or less lengthy) non-expressive one [47]. In this way, the observer who will see this two-segment block can get used to the encoder's facial features. However, a researcher concerned only with the most expressive parts can rely on the timelines. These allow for easy emotion segmentation of the recordings. Thanks to the timelines' precision, one can extract those excerpts of EFEs in which one is particularly interested. For instance, the end user can extract the frame that obtained the maximum amount of interdecoder agreement. Likewise, because the boundaries between the different emotional displays within a video are easy to spot, one can focus solely on the segment that reliably expresses a single target emotion.

5. CONCLUSION

Researchers need methods to record and to analyze human affective facial behaviors exhibiting dynamic and spontaneous EFEs accompanied by the encoder's affective states and decoders' annotations regarding their timecourse (see Zeng et al., 2009, for a review). For researchers working on emotion communication, or on the production and/or the perception and recognition of emotional facial expressions or working at the development of artificial systems designed to encode or decode human emotional signals, DynEmo is a comprehensive resource of filmed affective facial behaviors which provides a substantial publicly available corpus of validated dynamic and natural facial expressions of pervasive affective states (in line with Sneddon et al.,

2012) . It supplies 233 (Set 1) and 125 (Set 2) EFE recordings performed by a wide range of ordinary people, from young to older adults of both genders. Set 1 and Set 2 are labeled with the expresser's affective state (encoder's emotional self-report). The Set 2 is also provided with observers' dynamic ratings (decoder's emotional recognition timelines). All methodological, contextual, and other elements are at the disposal of the scientific community.

DynEmo is useful with respect to the distinction between production versus perception of affective behaviors. Firstly, regarding production, DynEmo provides genuine facial expressions — or first-order displays — exhibited in the course of a given eliciting episode. Moreover, the conditions that influence or cause affective behavior, whether internal or external to the encoder (her/his affective state, current situation, etc.), are known to the user. Secondly, data on self-reported emotions (Index Card) allows the user to study the extent to which an EFE reflects an emotion experienced by the encoder. For instance, data authorizes one to compare EFEs exhibited by encoders self-reporting high intensity of a given affective state to the ones of encoders self-reporting low intensity. Of course, people vary in their self-report of emotional experiences and self-report are probably an inaccurate reflection of the internal emotional world of the individual, but there is currently no external objective measure of the internal feelings of an individual. Irrespective of this drawback, emotional self-report data allows one to study the concordance/discrepancy between what is experienced and what is displayed. The analysis of natural and dynamic faces could reveal very different patterns of action units than the prototypical forms already known. Regarding perception/recognition, the dynamic aspects of the EFE recordings in DynEmo and their relationship to the observers' interpretations (i.e., from a decoding point of view) can be explored via the timelines. Such synchronized measures of encoding and decoding activities allows for a moment-by-moment analysis that simultaneously considers the encoder's facial changes and the observer's answers (independently of what is experienced by the encoder). Indeed, for each video emotional expression timelines instantly signal when a given affective state is considered to be displayed on the face. Time-lines thus allow the examination of main facial characteristics of EFEs, such as progression, intensity or emotional typology. For instance, one can analyze the time-course progression of different EFEs: their quick or slow on-sets and their quick or slow off-sets. One can also explore the intensity when looking the judges' agreements indicating if the displayed emotion is obvious or subtle. One can, finally, consider the emotional typology by examining the periods where observers decode different emotions. Emotional typology analysis illustrates if the expression is consensual or not, that is, when a strong or a weak consensus exists between judges regarding what is displayed by the face (distinctive /ambiguous expression). Emotional typology is particularly relevant, among others, for one interested in idiosyncratic characteristics of expressers.

Finally, the relation between what causes the face to move (production) and what an observer thinks causes the face to move (perception) is an empirical question — even if it has usually been an implicit assumption. Indeed, most researchers are interested by both aspects: what the facial behavior is like and what it signifies to the observer? Emotional recognition accuracy has widely been studied with posed and static stimuli. Mismatches between emotion displays and recognition performances are also of importance regarding the study of misinterpretations during communication processes as much as for researchers in the field of affective computing. The use of natural EFEs is usually considered as problematic regarding the question of accuracy. When using prototypical EFEs as the JACFEE set [48], the recognition of the emotion is estimated against the "accurate" expression displayed by the encoder. In the case of natural EFEs, the emotion to be recognized must be the one self-reported by the encoder. Indeed, using a natural EFE without knowing objectively what affective state is being experienced by the encoder is meaningless for one interested in recognition accuracy. Since researchers are most often interested in both aspects of production/perception, the DynEmo database is interesting for those who look at what is communicated by behavior and/or how it is interpreted by others. DynEmo timelines allow one to compare how well an emotion is communicated to observers.

DynEmo adds to the range of natural affect resources. It is freely accessible at the address given below and is powered by the Marvelig platform (LIG laboratory). To avoid circulation of DynEmo EFE videos not related to research usage, a consent paper must be signed by scientists to obtain a login on the site and to download the data from https://dynemo.upmf-grenoble.fr/.

ACKNOWLEDGEMENTS

REFERENCES

[1]   B. Schuller, M. Valstar, F. Eyben, R. Cowie, and M. Pantic, "AVEC 2012: the continuous audio/visual emotion challenge," in Proceedings of the 14th ACM international conference on Multimodal interaction, 2012, pp. 449-456.

[2]   Y. Tong, W. Liao, and Q. Ji, "Facial action unit recognition by exploiting their dynamic and semantic relationships," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, pp. 1683-1699, 2007.

[3]   Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, pp. 39-58, 2009.

[4]   R. Gross, "Face Databases," in Handbook of Face Recognition, ed New York, USA: Springer 2005, pp. 301-327.

[5]   C. Anitha, B. A. Venkatesha, and B. S. Adiga, "A survey on facial expression databases," International Journal of Engineering Science and Technology, vol. 2, pp. 5158-5174, 2010.

[6]   R. A. Calvo and S. K. D'Mello, New perspectives on affect and learning technologies vol. 3. New York, USA: Springer, 2011.

[7]   K. R. Scherer, T. Banziger, and E. B. Roesch, Blueprint for affective computing: A sourcebook. Oxford, UK: Oxford University Press, 2010.

[8]   P. Petta, C. Pelachaud, and R. Cowie, Emotion-oriented systems: the HUMAINE handbook. New York, USA: Springer, 2011.

[9]   R. B. Grossman and J. Kegl, "Moving faces: Categorization of dynamic facial expressions in american sign language by deaf and hearing participants," Journal of Nonverbal Behavior, vol. 31, pp. 23-38, 2007.

[10]   B. Gepner, C. Deruelle, and S. Grynfeltt, "Motion and emotion: A novel approach to the study of face processing by young autistic children," Journal of Autism and Developmental Disorders, vol. 31, pp. 37-45, 2001.

[11]   T. Wehrle, S. Kaiser, S. Schmidt, and K. R. Scherer, "Studying the dynamics of emotional expression using synthesized facial muscle movements," Journal of Personality and Social Psychology, vol. 78, pp. 105-119, 2000.

[12]   W. Sato and S. Yoshikawa, "The dynamic aspects of emotional facial expressions," Cognition and Emotion, vol. 18, pp. 701-710, 2004.

[13]   W. Sato and S. Yoshikawa, "Spontaneous facial mimicry in response to dynamic facial expressions," Cognition, vol. 104, pp. 1-18, 2007.

[14]   S. Yoshikawa and W. Sato, "Enhanced perceptual, emotional, and motor processing in response to dynamic facial expressions of emotion," Japanese Psychological Research, vol. 48, pp. 213-222, 2006.

[15]   J. F. Cohn and K. S. Schmidt, "The timing of facial motion in posed and spontaneous smiles," International Journal of Wavelets, vol. 2, pp. 1-12, 2004.

[16]   U. Hess and R. E. Kleck, "Differentiating emotion elicited and deliberate emotional facial expressions," European Journal of Social Psychology, vol. 20, pp. 369-385, 1990.

[17]   Z. Ambadar, J. W. Schooler, and J. F. Cohn, "Deciphering the enigmatic face," Psychological Science, vol. 16, p. 403, 2005.

[18] E. Bould and N. Morris, "Role of motion signals in recognizing subtle facial expressions of emotion," British Journal of Psychology, vol. 99, pp. 167-189, 2008.

[19] C. D. Kilts, G. Egan, D. A. Gideon, T. D. Ely, and J. M. Hoffman, "Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions," Neuroimage, vol. 18, pp. 156-168, 2003.

[20] R. Adolphs, "Recognizing emotion from facial expressions: Psychological and neurological mechanisms," Behavioral and Cognitive Neuroscience Reviews, vol. 1, pp. 21-62, 2002.

[21] A. J. Calder, "Categorical perception of morphed facial expressions," Visual Cognition, vol. 3, pp. 81-118, 1996.

[22] M. F. Valstar and M. Pantic, "Fully Automatic Recognition of the Temporal Phases of Facial Actions," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 42, pp. 28-43, 2012.

[23] A. Tcherkassof, T. Bollon, M. Dubois, P. Pansu, and J. M. Adam, "Facial expressions of emotions: A methodological contribution to the study of spontaneous and dynamic emotional faces," European Journal of Social Psychology, vol. 37, pp. 1325-1345, 2007.

[24] S. Wang, Z. Liu, Z. Wang, G. Wu, P. Shen, S. He, and X. Wang, "Analyses of a Multi-modal spontaneous facial expression database," IEEE Transactions on Affective Computing, vol. 3, pp. 1-14, 2012.

[25] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," presented at the International Conference on Automatic Face and Gesture Recognition, Grenoble, France, 2000.

[26] P. Ekman and E. L. Rosenberg, What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). New York, USA: Oxford University Press, 1997.

[27] M. F. Valstar, M. Pantic, Z. Ambadar, and J. F. Cohn, "Spontaneous vs. posed facial behavior: automatic analysis of brow actions," presented at the International Conference on Multimodal Interfaces (ICMI), Banff, Canada, 2006.

[28] W. G. Dopson, B. E. Beckwith, D. M. Tucker, and P. C. Bullard-Bates, "Asymmetry of facial expression in spontaneous emotion," Cortex, vol. 20, pp. 243-251, 1984.

[29] M. G. Frank, P. Ekman, and W. V. Friesen, "Behavioral markers and recognizability of the smile of enjoyment," Journal of Personality and Social Psychology, vol. 64, p. 83, 1993.

[30] M. F. Valstar, H. Gunes, and M. Pantic, "How to distinguish posed from spontaneous smiles using geometric features," presented at the International Conference on Multimodal Interfaces (ICMI), Nagoya, Japan, 2007.

[31] K. R. Scherer and H. Ellgring, "Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal?," Emotion, vol. 7, pp. 113-130, 2007.

[32] E. Douglas-Cowie, R. Cowie, I. Sneddon, C. Cox, O. Lowry, M. McRorie, J. C. Martin, L. Devillers, S. Abrilian, and A. Batliner, "The humaine database: Addressing the collection and annotation of naturalistic and induced emotional data," presented at the Affective Computing and Intelligent Interaction Conference, Berlin, Germany, 2007.

[33] D. E. Back, T. R. Jordan, and S. M. Thomas, "The recognition of mental states from dynamic and static facial expressions," Visual Cognition, vol. 17, pp. 1271-1286, 2009.

[34] J. M. Fernandez-Dols, P. Carrera, K. A. Barchard, and M. Gacitua, "False recognition of facial expressions of emotion: Causes and implications," Emotion, vol. 8, pp. 530-539, 2008.

[35] H. A. Elfenbein and N. Ambady, "On the universality and cultural specificity of emotion recognition: a meta-analysis," Psychological bulletin, vol. 128, pp. 203-235, 2002.

[36] J. M. Fernandez-Dols, J. A. Russell, and P. Carrera, "Are facial displays social? Situational influences in the attribution of emotion to facial expressions," The Spanish Journal of Psychology, vol. 5, pp. 119-124, 2002.

[37] I. Sneddon, M. McRorie, G. McKeown, and J. Hanratty, "The Belfast Induced Natural Emotion Database," IEEE Transactions on Affective Computing, vol. 3, pp. 32-41 2012.

[38] G. McKeown, M. F. Valstar, R. Cowie, and M. Pantic, "The SEMAINE corpus of emotionally coloured character interactions," presented at the International Conference on Multimedia and Expo (ICME), Suntec City, Singapore, 2010.

[39] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," Affective Computing, IEEE Transactions on, vol. 3, pp. 42-55, 2012.

[40] J. A. Russell, "A circumplex model of affect," Journal of Personality and Social Psychology, vol. 39, pp. 1161-1178, 1980.

[41] N. H. Frijda and A. Tcherkassof, "Facial expressions as modes of action readiness," in The psychology of facial expression, James A. Russell and J. M. Fernández-Dols, Eds., ed Cambridge, UK: Cambridge University Press, 1997, pp. 78-102.

[42] B. Meillon, A. Tcherkassof, N. Mandran, J. M. Adam, M. Dubois, D. Dupré, A. M. Benoit, A. Guérin-Dugué, and A. Caplier, "DYNEMO: A Corpus of dynamic and spontaneous emotional facial expressions," in LREC, Malte, Malte, 2010.

[43] O. Langner, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, S. T. Hawk, and A. Van Knippenberg, "Presentation and validation of the Radboud Faces Database," Cognition & Emotion, vol. 24, pp. 1377-1388, 2010.

[44] J. L. Tracy, R. W. Robins, and R. A. Schriber, "Development of a FACS-verified set of basic and self-conscious emotion expressions," Emotion, vol. 9, pp. 554-559, 2009.

[45] J. M. Bland and D. G. Altman, "A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement," Computers in biology and medicine, vol. 20, pp. 337-340, 1990.

[46] K. O. McGraw and S. P. Wong, "Forming inferences about some intraclass correlation coefficients," Psychological methods, vol. 1, pp. 30-46, 1996.

[47] J. Van der Schalk, S. T. Hawk, A. H. Fischer, and B. Doosje, "Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES)," Emotion, vol. 11, pp. 907-920, 2011.

[48] D. Matsumoto and P. Ekman, "Japanese and Caucasian facial expressions of emotion and neutral faces (JACFEE and JACNeuF)," H. I. Laboratory, Ed., ed. San Francisco, USA, 1988.

APPENDIX 1.

Self-report index card available for each encoder's recording

APPENDIX 2.

Encoders' affective self-reports

To evaluate encoders' self-report on dimensional, action tendency and emotional scales (respectively Table A, Table B and Table C, mean scores for each elicitating task were compared with mean scores for the neutral task. Mean comparisons were calculated thanks to independent t test with Bonferroni corrections.

The first scale concerns the dimensional items (Table A). Encoders reported feeling significantly displeasure during negative task 5, 6 and 8 tasks, as compared to encoders during the neutral task. No difference between the tasks is observed as regard to *Arousal* and *Focus* items. This could be explained by the very high averages suggesting a great involvement of encoders in achieving their task (ceiling effect). Finally, encoders reported significantly less comfort during negative tasks 6 and 8 as compared to the neutral task.

Table A. Means and standard deviance for the dimensional scales. Means in bold indicate a significant difference between the elicitating task and the neutral one.

Self-report	Tasks	P1	P2	P3	P4	N5	N6	N7	N8	N9	N10	NE
Valence	M	5.05	4.23	5.22	5.05	**3.15***	**2.55***	4.45	**2.63***	4.10	4.60	4.82
	SD	1.2	1.3	0.8	0.9	1.4	0.9	1.0	1.2	1.6	1.4	0.9
Arousal	M	4.59	4.10	4.85	4.84	4.30	3.31	4.19	3.57	4.65	4.90	4.27
	SD	0.9	1.3	0.9	1.0	1.2	0.9	1.2	1.1	1.1	1.1	1.4
Focus	M	5.03	4.60	5.46	4.71	4.80	4.86	4.38	4.63	4.60	5.10	4.64
	SD	1.0	1.1	0.7	0.9	1.0	0.9	1.3	1.2	1.3	1.0	1.3
Confort	M	5.08	4.48	5.22	5.26	3.53	**2.93***	4.33	**2.91***	4.30	4.50	4.91
	SD	1.2	1.4	1.1	1.0	1.6	1.2	1.4	1.4	1.5	1.5	0.9

*p < .05

On the whole, few action readiness modes differentiate the eliciting tasks from the neutral task (Table B). Negative task 8 stands out with the items *To protect oneself*, *To have anything to do with* and *To be tensed* as compared to the neutral task. The negative task 6 stands out for the *Not to exist* and *To want to cry* items. Negative task 3 and 4 are significantly lower than the neutral task for the item *To feel inhibited*. Finally, encoders in the negative tasks 5, 6 and 7 rated significantly less the item *To want to laugh* as compared to the neutral task.

Table B. Mean and standard deviance for the action tendency scale. Means in bold indicate a significant difference between the elicitating task and the neutral one.

Self-report		P1	P2	P3	P4	N5	N6	N7	N8	N9	N10	NE
To approach	M	2.51	2.75	2.49	2.42	2.60	3.52	2.17	1.91	2.55	3.20	3.09
	SD	1.2	1.5	1.6	1.4	1.5	1.4	1.3	1.1	2.0	2.0	1.6
To be with	M	2.46	2.33	2.07	2.08	2.78	3.38	2.12	1.97	1.95	3.00	3.09
	SD	1.3	1.4	1.4	1.3	1.6	1.5	1.2	1.1	1.6	2.1	1.6
To protect	M	2.15	2.33	1.44	1.42	2.58	2.52	1.69	**3.20***	1.70	1.50	1.36
	SD	1.5	1.5	1.0	0.8	1.6	1.3	1.2	1.6	1.6	1.3	0.5
To stay away	M	2.10	1.83	1.32	1.53	2.35	1.93	1.67	3.00	1.50	2.10	2.18
	SD	1.5	1.1	0.8	0.8	1.5	1.1	1.1	1.6	1.0	1.7	1.5
To pay attention	M	2.23	2.48	1.27	1.68	2.40	2.12	1.88	3.51	1.30	1.40	2.18
	SD	1.5	1.5	0.7	1.0	1.3	1.1	1.2	1.6	0.7	0.7	1.0
To keep at distance	M	2.92	3.48	2.83	3.50	3.08	4.19	3.05	3.54	3.30	3.10	3.73
	SD	1.8	1.8	1.9	1.8	1.7	1.5	1.8	1.5	2.0	2.0	1.4
To reject	M	2.13	1.73	1.20	1.50	2.75	1.83	1.76	**4.51***	1.70	1.50	1.45
	SD	1.5	1.0	0.6	0.7	1.5	1.0	1.2	1.5	1.4	0.8	0.7
Disinterest	M	1.64	1.93	1.51	1.84	2.05	1.57	2.62	2.23	2.05	2.00	1.73
	SD	1.0	1.1	1.1	1.1	1.3	0.8	1.7	1.3	1.5	1.9	1.0
Not to exist	M	2.62	2.05	1.61	1.92	3.48	**5.17***	2.69	3.60	2.30	1.70	3.27
	SD	1.5	1.3	1.2	1.3	1.9	1.2	1.7	1.8	1.8	1.1	1.9
To boil inside	M	1.97	1.63	1.32	1.61	2.45	3.98	2.05	3.34	2.90	1.40	2.82
	SD	1.2	0.9	0.7	0.8	1.6	1.5	1.5	1.7	2.0	0.5	2.0
To oppose	M	1.87	1.63	**1.24***	**1.42***	2.30	3.93	1.95	3.40	2.00	1.60	3.00
	SD	1.3	0.9	0.7	0.6	1.6	1.4	1.2	1.7	1.2	1.0	1.7
To go against sth	M	2.46	1.93	1.66	1.63	2.43	3.52	2.19	2.91	2.05	1.80	2.91
	SD	1.6	1.2	1.2	1.0	1.4	1.3	1.3	1.3	1.4	0.9	1.8
To interrupt	M	2.33	2.45	1.41	1.71	2.60	3.00	1.93	3.80	2.20	1.60	2.27
	SD	1.5	1.6	1.0	1.1	1.7	1.5	1.4	1.7	1.5	1.1	1.6
To merge with	M	1.85	1.93	1.63	1.76	1.95	2.64	1.57	2.03	1.80	2.70	2.55
	SD	1.1	1.2	1.1	1.1	1.2	1.4	1.0	1.3	1.4	1.6	1.8

Not concentrate	M	1.85	2.20	1.12	1.63	2.43	2.90	2.36	3.40	1.80	1.70	2.09
	SD	1.0	1.2	0.3	0.8	1.6	1.5	1.7	1.5	1.0	0.9	1.3
To stand above	M	3.10	3.50	2.88	2.79	2.78	2.48	2.38	3.31	3.30	3.70	3.55
	SD	1.7	1.7	1.7	1.6	1.5	1.5	1.6	1.6	1.9	1.9	1.0
To help someone	M	2.97	2.30	2.10	2.05	3.50	4.33	1.83	2.71	2.40	1.80	3.36
	SD	1.8	1.5	1.5	1.4	1.7	1.6	1.3	1.6	1.9	1.2	1.8
To disappear	M	1.72	1.73	1.12	1.29	1.75	1.93	1.43	2.03	1.40	1.60	1.55
	SD	1.0	1.2	0.4	0.5	1.3	1.4	0.9	1.2	1.1	1.0	0.9
To feel inhibited	M	1.87	1.90	1.41*	1.63*	2.10	2.24	2.21	3.06	2.20	1.50	3.27
	SD	1.2	1.0	0.7	0.9	1.5	1.4	1.6	1.6	1.7	1.0	1.8
To not want to	M	1.97	1.88	1.37	1.58	1.45	1.74	1.81	1.74	1.60	1.90	2.55
	SD	1.6	1.3	0.8	1.1	0.9	1.3	1.3	1.0	1.1	1.2	1.5
To be apathetic	M	1.64	1.55	1.07	1.37	1.53	1.69	1.83	1.74	1.15	1.50	1.91
	SD	1.1	0.7	0.3	0.6	1.0	0.9	1.2	0.8	0.4	1.0	1.0
To give up	M	1.49	1.40	1.07	1.29	1.48	1.60	1.52	1.83	1.70	1.10	1.55
	SD	0.8	0.6	0.3	0.5	0.8	0.9	0.8	1.1	1.3	0.3	0.8
To shut oneself	M	1.49	1.68	1.15	1.42	1.58	1.57	1.67	2.31	1.45	1.20	1.91
	SD	0.8	1.0	0.4	0.6	1.1	0.9	1.3	1.3	0.9	0.4	1.4
To be helpless	M	1.67	2.15	1.41	1.71	2.55	3.10	2.24	2.51	1.90	1.20	1.91
	SD	1.0	1.4	0.9	1.0	1.8	1.5	1.4	1.4	1.4	0.4	0.9
To want to cry	M	1.28	1.30	1.07	1.26	1.38	2.90*	1.17	1.63	1.25	1.10	1.45
	SD	0.5	0.5	0.3	0.5	1.0	1.4	0.4	1.1	0.9	0.3	0.8
To be excited	M	1.59	1.98	1.54	1.55	1.80	1.69	1.74	3.06	2.15	1.70	2.18
	SD	0.8	1.2	0.9	0.7	1.3	0.8	1.3	1.3	1.7	0.9	1.7
To be exuberant	M	1.54	1.60	1.44	1.37	1.30	1.40	1.38	1.97	2.05	1.20	2.09
	SD	0.9	1.0	0.8	0.6	0.6	0.7	0.9	0.9	1.7	0.4	1.6
To want to laugh	M	4.67	2.33	2.71	2.92	1.53*	1.26*	1.60*	2.60	4.50	3.20	3.91
	SD	1.6	1.5	1.9	1.8	1.1	0.6	1.0	1.8	1.7	2.1	1.8
To feel at rest	M	1.85	1.88	1.71	1.82	2.65	2.74	1.79	4.11	2.25	2.30	2.36
	SD	1.2	1.1	1.2	1.2	1.5	1.4	1.0	1.7	1.7	1.3	1.1
To be tensed	M	4.51	3.73	3.76	3.89	2.93	2.29	3.71	2.23*	3.95	4.20	3.18
	SD	1.4	1.6	1.7	1.6	1.4	1.4	1.5	1.6	1.8	1.5	1.5
To do an effort	M	2.26	3.48	3.12	2.87	2.70	3.00	2.98	2.66	3.55	1.60	3.64
	SD	1.4	1.8	1.8	1.7	1.5	1.4	1.7	1.5	1.7	1.1	1.2
To be open	M	4.62	4.85	4.56	4.79	4.45	4.76	4.60	4.09	3.90	5.10	4.91
	SD	1.4	1.2	1.5	1.5	1.4	1.2	1.6	1.5	1.9	1.1	1.2
To embody	M	2.15	2.48	1.71	1.76	1.85	2.05	2.12	1.97	1.95	1.60	1.82
	SD	1.4	1.6	1.2	1.1	1.2	1.1	1.5	1.2	1.6	1.1	1.4
To be tensed	M	1.82	2.18	1.46	1.71	2.30	2.86	1.93	3.37	2.20	1.20	2.73
	SD	0.9	1.1	1.0	1.0	1.5	1.3	1.4	1.7	1.6	0.6	1.4
To be still	M	2.18	2.58	1.71	2.24	2.20	2.43	2.95	2.51	1.65	1.80	2.00
	SD	1.3	1.7	1.3	1.6	1.5	1.5	1.9	1.4	1.3	1.3	1.5

*p < .05

Regarding the emotional labels scales, all encoders reported a high Curiosity experience but no means differs significantly from the neutral task (Table C). These high mean show the encoders' interest for the experiment in general. Three other labels do not significantly differ from the neutral task: *Astonishment*, *Disappointment* and *Humiliation*. Otherwise, encoders of positive task 3 reported to feel less boredom than those running the neutral task. Encoders of negative task 5 and 8 were significantly more disgusted than those running the neutral one. Encoders of negative task 6 and 8 were significantly more frightened, encoders of negative tasks 5 and 6 were significantly less amused, encoders of the negative task 6 felt significantly more shame and sadness that encoders of the neutral task. Finally, encoders of negative tasks 1, 4, 5, 6, 8, 9 and 10 reported *Satisfaction* significantly less than encoders of the neutral task and encoders of positive tasks 1, 2, 3, 4, negative tasks 5 and 7 felt significantly less *Annoyance* than encoders of the neutral task.

Table C. Mean and standard deviance for the emotional labels scales. Means in bold indicate a significant difference between the eliciting task and the neutral one.

Self-report	Tasks	P1	P2	P3	P4	N 5	N6	N7	N 8	N 9	N10	NE
Curiosity	M	5.28	5.38	5.27	5.32	5.28	3.52	5.36	4.83	5.20	4.60	4.73
	SD	0.8	1.0	1.2	0.9	0.8	1.8	0.9	1.0	1.2	0.8	1.2
Boredom	M	1.38	1.35	1.15*	1.47	1.43	1.43	2.10	1.69	1.65	1.30	2.36
	SD	0.7	0.6	0.4	0.9	0.9	0.8	1.7	1.1	0.9	0.7	1.2
Disgust	M	1.62	1.28	1.10	1.21	3.55*	2.45	1.33	3.34*	1.40	1.40	2.00
	SD	1.1	0.6	0.3	0.4	1.8	1.5	0.7	1.7	0.8	1.0	1.3
Fright	M	1.51	1.53	1.15	1.21	2.90	3.02*	1.36	4.80*	1.40	1.30	1.64
	SD	1.0	0.9	0.4	0.4	1.7	1.7	0.7	1.4	1.1	0.7	0.8

Amusement	M	4.64	3.15	3.44	3.61	**1.28**	**1.64**	1.86	2.14	3.55	3.50	3.36
	SD	1.2	1.4	1.6	1.4	0.7	1.0	1.4	1.3	2.0	1.1	1.4
Shame	M	1.21	1.20	1.07	1.26	1.23	**3.81***	1.26	1.54	1.55	2.30	1.82
	SD	0.5	0.4	0.3	0.6	0.5	1.6	0.6	1.0	0.9	1.3	1.0
Satisfaction	M	**2.00***	2.40	2.90	**2.00***	**1.23***	**1.40***	2.31	**1.69***	**1.85***	**1.50***	3.55
	SD	1.3	1.6	1.6	1.4	0.6	0.6	1.6	1.2	1.6	1.1	1.8
Astonish.	M	2.46	3.23	2.34	2.29	3.98	2.60	3.29	4.46	2.40	2.60	2.73
	SD	1.5	1.8	1.7	1.6	1.9	1.6	1.9	1.7	1.6	1.2	1.7
Moved	M	2.36	2.18	1.46	1.53	3.20	**5.05***	1.81	3.11	1.45	2.50	2.82
	SD	1.5	1.5	1.1	0.7	1.7	1.1	1.2	1.7	1.1	1.4	1.5
Annoyance	M	**1.38***	**1.43***	**1.12***	**1.21***	**1.40***	2.64	**1.50***	2.66	2.40	1.30	2.82
	SD	0.7	0.9	0.4	0.4	0.9	1.6	0.8	1.6	1.8	0.9	1.9
Disappoint.	M	1.23	1.25	1.10	1.34	1.10	1.60	1.29	1.83	1.95	1.30	1.91
	SD	0.5	0.5	0.3	0.6	0.3	0.9	0.9	1.2	1.6	0.9	1.4
Humiliation	M	1.28	1.68	1.34	1.68	1.40	2.21	1.52	2.54	1.55	1.80	2.55
	SD	0.6	1.1	0.8	1.1	0.9	1.6	1.2	1.5	0.9	1.8	1.5

$*p < .05$

APPENDIX 3.

Emotion identification scores for decoded recordings (Set 2): the frame during which a maximum of decoders have attributed the same emotion label (MA for Maximum agreement) and the Intraclass Correlation (ICC) for the recording.

Video	N	ICC	MA	frame for MA (in 0.1 sec)	Label for MA	Video	N	ICC	MA	frame for MA (in 0.1 sec)	Label for MA
DVD02_5	24	.75	0.83	437	Disgust	DVD29_1	24	.91	0.83	74	Astonishment
DVD03_1	22	.90	0.59	147	Astonishment	DVD29_4	22	.70	0.59	196	Fright
DVD03_5	21	.71	0.43	88	Astonishment	DVD29_5	26	.53	0.38	1407	Boredom
DVD04_1	23	.88	0.35	1742	Boredom	DVD30_3	22	.85	0.68	494	Amusement
DVD04_2	18	.76	0.39	1226	Boredom	DVD31_1	20	.95	0.95	1603	Amusement
DVD04_3	23	.94	0.70	559	Amusement	DVD31_2	19	.91	1.00	1135	Amusement
DVD05_3	26	.73	0.65	516	Astonishment	DVD31_4	25	.76	0.24	755	Boredom
DVD06_2	21	.65	0.14	1223	Astonishment	DVD36_1	20	.14	0.15	340	Astonishment
DVD06_3	24	.87	0.63	1373	Astonishment	DVD37_2	22	.90	0.86	114	Astonishment
DVD06_4	26	.86	0.62	1012	Amusement	DVD38_5	22	.87	0.68	825	Boredom
DVD07_1	22	.90	0.32	185	Amusement	DVD39_4	22	.80	0.59	382	Astonishment
DVD07_3	21	.90	0.38	90	Boredom	DVD44_3	23	.34	0.30	37	Boredom
DVD08_4	22	.65	0.41	87	Curiosity	DVD45_4	21	.86	0.95	401	Amusement
DVD09_1	25	.78	0.60	398	Astonishment	DVD49_5	23	.92	0.74	90	Astonishment
DVD09_3	24	.93	0.83	629	Amusement	DVD49_6	25	.94	0.84	624	Amusement
DVD09_5	22	.83	0.77	1608	Amusement	DVD51_5	22	.85	0.68	107	Astonishment
DVD10_1	21	.80	0.43	1933	Amusement	DVD52_1	24	.83	0.58	373	Amusement
DVD10_3	22	.85	0.45	1266	Curiosity	DVD53_4	26	.93	0.46	74	Amusement
DVD11_1	20	.89	0.55	138	Astonishment	DVD53_5	20	.75	0.65	403	Amusement
DVD11_2	25	.96	0.72	866	Amusement	DVD54_1	20	.76	0.25	352	Curiosity
DVD11_3	20	.90	0.55	138	Astonishment	DVD60_2	23	.74	0.26	275	Moved
DVD11_5	21	.93	0.90	55	Astonishment	DVD63_1	19	.81	0.53	306	Curiosity
DVD12_1	26	.93	0.88	84	Astonishment	DVD64_3	21	.87	0.71	83	Astonishment
DVD12_3	22	.42	0.27	1213	Boredom	DVD68_1	25	.77	0.44	356	Astonishment
DVD12_6	20	.95	0.40	733	Curiosity	DVD69_3	22	.94	0.64	763	Amusement
DVD13_6	21	.66	0.33	1098	Amusement	DVD69_4	23	.90	0.78	123	Astonishment
DVD14_1	20	.95	0.95	504	Amusement	DVD70_3	23	.79	0.17	55	No Emotion
DVD14_2	20	.90	0.70	587	Amusement	DVD70_6	22	.01	0.14	1691	Boredom
DVD14_3	21	.80	0.48	1406	Amusement	DVD71_2	23	.45	0.26	1450	Boredom
DVD14_4	23	.76	0.35	1313	Astonishment	DVD72_3	24	.90	0.67	741	Boredom

DVD15_1	21	.89	0.76	447	Astonishment	DVD73_2	23	.67	0.52	280	Boredom
DVD18_5	20	.90	0.80	522	Astonishment	DVD74_3	26	.87	0.77	1274	Amusement
DVD20_3	23	.88	0.22	3259	No Emotion	DVD74_4	25	.89	0.96	622	Amusement
DVD21_1	25	.93	0.88	615	Amusement	DVD75_4	26	.81	0.27	129	Astonishment
DVD21_2	22	.90	0.64	154	Fright	DVD76_1	21	.75	0.33	1469	Disappointment
DVD22_2	21	.88	0.62	947	Amusement	DVD76_2	25	.68	0.32	1003	Disappointment
DVD23_1	22	.85	0.82	1653	Disgust	DVD76_3	23	.70	0.39	847	Disappointment
DVD23_3	22	.93	0.77	63	Astonishment	DVD77_4	19	.92	0.58	1101	Amusement
DVD24_1	17	.60	0.29	1749	Disgust	DVD78_3	24	.89	0.46	718	Amusement
DVD24_2	20	.84	0.65	1695	Curiosity	DVD78_4	23	.77	0.30	291	Boredom
DVD24_3	23	.83	0.39	146	Fright	DVD79_5	23	.91	0.87	705	Amusement
DVD26_1	19	.82	0.68	412	Disgust	DVD80_1	24	.45	0.25	1678	Boredom
DVD27_4	20	.47	0.20	956	Fright	DVD80_2	21	.37	0.19	495	Boredom
DVD28_2	23	.93	0.39	487	Astonishment	DVD80_3	21	.85	0.95	978	Amusement
DVD28_5	20	.94	1.00	585	Amusement	DVD80_4	21	.94	0.90	1050	Amusement
DVD28_6	22	.77	0.23	230	Boredom						