Running head : ASSESSING EMOTIONAL FILMS

Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers.

Alexandre Schaefer[1], Frédéric Nils[2], Xavier Sanchez[3], and Pierre Philippot[2]

[1]University of Leeds, UK, [2]University of Louvain, Louvain-La-Neuve, Belgium, [3]University of Chester, UK .

Authors' notes

Correspondence regarding this manuscript should be addressed to Alexandre Schaefer who is at the University of Leeds, Institute of Psychological Sciences, Leeds LS2 9JT, UK. Email: a.schaefer@leeds.ac.uk or Pierre Philippot, Université Catholique de Louvain, Faculté de Psychologie, 10, Place du Cardinal Mercier, 1348-Louvain-La-Neuve, Belgium. Email : Pierre.Philippot@uclouvain.be. The database of stimuli can be found at this URL: http://nemo.psp.ucl.ac.be/FilmStim/. The writing of this article has been facilitated by grants from the “Fonds National de la Recherche Scientifique” (FNRS) 8.4510.99 and 8.4510.00. The authors thank Deanna Barch, Todd Braver and Jeremy Gray for their helpful comments on an earlier draft of this manuscript. The authors also wish to thank Sylvie De Coninck, Cécile Misonne and Olivier Beauclercq for their contribution to the data collection, Pierre Mahau and Richard Robert for their help in recording and editing the film excerpts, and Adam Rickart for helping with the data analysis.

Abstract

Using emotional film clips is one of the most popular and effective methods of emotion elicitation. The main goal of the present study was to develop and test the effectiveness of a new and comprehensive set of emotional film excerpts. Fifty film experts were asked to remember specific film scenes that elicited fear, anger, sadness, disgust, amusement, tenderness, as well as emotionally neutral scenes. For each emotion, the 10 most frequently mentioned scenes were selected and cut into film clips. Next, 364 participants viewed the film clips in individual laboratory sessions and rated each film on multiple dimensions. Results showed that the film clips were effective with regard to several criteria such as emotional discreteness, arousal, positive and negative affect. Finally, ranking scores were computed for 24 classification criteria: Subjective arousal, positive and negative affect (derived from the PANAS, Watson & Tellegen, 1988), a positive and a negative affect scores derived from the Differential Emotions Scale (DES, Izard et al., 1974), six emotional discreteness scores (for anger, disgust, sadness, fear, amusement and tenderness), and 15 “mixed feelings” scores assessing the effectiveness of each film excerpt to produce blends of specific emotions. In addition, a number of emotionally neutral film clips were also validated. The database and editing instructions to construct the film clips were made freely available in a website.

Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers.

The increasing interest in the effects of emotion on cognitive, social and neural processes creates a constant need for efficient and reliable techniques of emotion elicitation. Many emotion elicitation techniques have been used so far: Exposure to emotional slides (Bradley & Lang, 2000; Schaefer et al., 2009), autobiographical recollection (Schaefer & Philippot, 2005), mental imagery (Schaefer et al., 2003; Vrana, Cuthbert & Lang, 1986), exposure to emotional film excerpts (Gray et al., 2005, 1995; Philippot, Schaefer & Herbette, 2003, Schaefer et al., 2006), Velten mood induction technique (Velten, 1969), facial feedback (Matsumoto, 1987), respiratory feedback (Philippot, Chapelle & Blairy, 2002), real-life techniques (Landis, 1924; Stemmler et al., 2001), and many other techniques.

Compared to other methods, exposure to emotional film excerpts has several advantages: First, it is one of easiest techniques to implement in a laboratory. Second, it has been widely observed that film excerpts can elicit strong subjective and physiological changes (e.g. Frazier, Strauss & Steinhauer, 2004; Gross, 1998; Palomba, Sarlo, Grilli et al., 2000). Third, the dynamic nature of film scenes provides an optimal artificial model of reality, without the ethical and practical problems of real-life techniques. Fourth, it seems to be the most powerful technique to elicit emotion in a laboratory: Westermann, Spies, Stahl & Hesse (1996) showed in a meta-analysis that among several other techniques, film clips were the most potent inducers of both positive and negative mood states.

Previous studies have already tested the reliability of film clips to elicit emotions, leading to the development of reliable sets of stimuli (e.g. Gross & Levenson, 1995; Hagemann et al., 1999; Mc Hugo et al, 1982; Philippot, 1993). The present validation study aimed to set up a new battery of film stimuli that expands the current choice of film stimuli and allows a more flexible selection of stimuli covering a larger array of potential research questions. Therefore, we developed and tested the effectiveness of what is to our knowledge the largest normative data base of emotional film scenes. The present article aims at

describing the development of this database, including the tests that attested its efficiency in eliciting emotions in a laboratory context.

What tasks need to be achieved by film stimuli in emotion research? Traditionally, it has been accepted that film stimuli must elicit differentiated emotional states such as anger or fear. This approach was linked to the "basic emotions" framework, i.e. the view that emotions are organized in a finite set of basic categories. For instance, Ekman (1984, 1992) suggested that there are 6 basic emotions: happiness, surprise, fear, sadness, anger and disgust. These views have led to the creation of sets of emotional film stimuli able to elicit differentiated experiential states corresponding to basic emotions. However, emotion research often addresses questions beyond emotional discreteness and hence needs stimuli validated on more diverse criteria. For instance, two films inducing differentiated feelings of sadness and happiness would not be useful for a study concerned with the effects of different levels of emotional intensity (or arousal) (e.g. low vs. moderate vs. high) on a cognitive task (e.g. Rottenstreich & Hsee, 2001). In fact, a significant proportion of emotion research uses parameters derived from a dimensional approach of emotions, in which emotional states are thought to be organized along generic continuous dimensions rather than discrete basic emotions (e.g. Russell, 1980; Plutchik, 1980; Davidson, 1992, 1993; Lang et al., 1993; Watson & Tellegen, 1988). Intensity (or arousal) and valence (i.e. the dissociation between positive an negative states) tend to be very popular dimensions used in emotion research (Bradley & Lang, 2001; Gray , 2001). Therefore, a battery of emotional stimuli needs to provide contents allowing the manipulation of these dimensions as well as the manipulation of basic emotions. Moreover, there is also the need for some researchers to elicit "blends" of emotions rather than pure, discrete states. An emerging trend of research has focused on how "mixed feelings" can arise, i.e. the simultaneous presence of more than one basic emotion (Schimmack, 2001; Hemenover & Schimmack, 2007). Therefore, an ideal database of emotional stimuli would also provide the possibility of selecting stimuli on the basis of their capacity to elicit specific "mixes" of emotional states.

In addition, a database of emotion stimuli needs to offer the possibility to tap a wide array of emotional experiences. Regarding this issue, there is a common bias in emotion research in favour of negative emotions, which have traditionally received more attention than positive emotions in the literature (Averill, 1982; Buck, 1999). This phenomenon is compounded by the fact that negative emotions might be more differentiated than positive emotions (Shaver et al., 1987). Positive emotions are generally presented under a generic label of *happiness*. More particularly, *attachment-related emotions* (e.g. social closeness, love, affection, tenderness, empathy) are generally underrepresented in emotion stimuli. These emotional states are thought to result from biological basic tendencies of *protection* and *contact* (Panksepp, 1998; Frijda, 1986), and are related to the concept of *affiliation* (Depue & Morrone-Strupinsky, In Press). Therefore, a comprehensive battery of emotional stimuli needs to include materials related to attachment-related emotions alongside other positive emotions.

In sum, emotion induction in an experimental setting might vary according to different parameters. The most often studied parameters are (a) the type of basic emotion induced—which implies the need to induce "pure" basic emotions, or alternatively, the need to induce blends of specific emotions (b) the intensity (or perceived arousal) of the feeling state induced, and (c) its valence. The ideal battery of emotion-eliciting film stimuli should thus offer to researchers the possibility of selecting, according to their research needs, stimuli validated on these different parameters. In addition, an ideal battery of emotional stimuli needs to include a wide choice of emotional experiences – including attachment-related emotions.

Overview

The main purpose of the present study was to set up a battery of emotion-eliciting film stimuli that would be as adaptable as possible to a large array of research questions. Therefore it is important to stress that it was not the goal of this study to investigate mechanisms of emotion *per se*, but instead to develop and assess the effectiveness of a

methodological tool for emotion research. The present project had then the following goals: (1) To select a large set of emotion-eliciting film scenes covering a wide range of emotional dimensions; (2) To include subsets of positive films that can elicit attachment-related emotions; (3) To assess the effectiveness of the film set along several criteria; (4) To provide a method of classification of films along multiple criteria, and identify subgroups of films with high scores on every criterion; (5) To make the data freely available on the Internet allowing a flexible choice of stimuli by potential users.

The present article will first describe how the film battery was created. Second, we will describe how the effectiveness of the films as emotional stimuli was tested. Specifically, we asked a large number of participants to watch the film excerpts in a laboratory setting. These emotional responses were assessed with three standard self-report measures: an extended version of the Differential Emotional Scale (DES, Izard et al, 1974), the PANAS scales (Watson & Tellegen, 1988) and a scale of subjective emotional arousal. The DES was used to assess discrete emotional states in order to check whether the film stimuli could trigger differentiated emotional feeling states consistent with a Basic emotions Approach (Ekman, 1984, 1992; Izard, 1991). The PANAS scales were used because they can assess two independent factors of positive and negative affect (Watson & Tellegen, 1988). This enabled us to assess the film set using a tool consistent with a dimensional approach of emotions (e.g. Russell, 1980; Plutchik, 1980; Davidson, 1992, 1993; Lang et al., 1993; Watson & Tellegen, 1988). Next we also measured arousal, which taps another important emotional dimension (Russell, 1980, Lang et al., 1993; Sonnemans & Frijda, 1994).

Method

Preliminary survey and excerpt selections

The first step in the present study was the selection of a relatively large number of film excerpts corresponding to seven *a priori* emotional categories: anger, sadness, fear, disgust, amusement, tenderness, and neutral state. The choice of the first four categories was motivated by the assumptions (a) that they cover the negative states most commonly

accepted as being “basic emotions” (Ekman, 1984, 1992; Izard, 1991), (b) that they are often studied in emotion research, and (c) that they refer to verbal labels easy to understand and commonly used in everyday language. Amusement and tenderness are not generally regarded as canonical basic emotions but they were chosen because they refer to two distinct positive categories evidenced in previous research on emotion categorization (Shaver et al., 1987). Further, and pragmatically, they can be easily evoked by films. In addition, tenderness is an attachment-related emotion (Hatfield & Rapson, 2000) therefore filling a gap in the field of emotion elicitation.

In order to provide a fairly large number of film clips, 10 films were selected per emotional category. These 70 film excerpts had to be selected from a large initial pool of possible film scenes. For this purpose, 50 experts (i.e. independent film rental store managers) were asked to fill in a brief questionnaire in which they had to recall, describe and rate for emotional arousal at least three film excerpts for each emotional category. Film rental store managers were chosen because they were likely to have a significant knowledge about films, which could contribute to create a large preliminary list of film scenes. A large initial set of 824 film excerpts was obtained with this method. For each emotional category, the 10 most frequently cited excerpts were chosen. If two or more films were quoted with the same frequency, the most intense film was chosen. The 70 selected excerpts were cut and recorded on VHS videotapes[1].

The duration of each film clip ranged from 1 to 7 min. Each clip was cut to make a coherent segment in order to maximize the emotional meaning of each clip. For instance, the excerpt of "American History X", in which a violent murder is shown, is cut just before the police sirens can be heard. The purpose was to prevent an attenuation of the meaning of injustice which is evoked by the scene. As participants were all French-speaking students, the spoken language of all excerpts was French (Dubbed versions were used for all non-French films).

Participants and Design

Three hundred and sixty-four undergraduate students (294 females and 70 males, mean age: 19.6, *SD* = 3.11) from a Belgian French-speaking university took part in the experiment for course credits. Participants were pseudo-randomly assigned to 7 groups, such that each group had an equal proportion of male participants. Each group had an average of 52 participants (max: 56 and min: 44). Within each group, 10 subgroups were created, differing only according to the order of presentation of the stimuli, to control for potential order effects. The order of presentation was also set in such a way that (a) two films targeting the same emotion were not shown consecutively, and (b) participants never watched more than two films of the same valence consecutively. A set of 10 different film excerpts was assigned to each group. Within each 10-films set, 7 film categories (anger, happiness, fear, disgust, sadness, tenderness and neutral) were represented. Consequently, three film categories were represented twice in each group. The repetition of film categories was fully counterbalanced across the seven groups. For subject-level analyses (in which participants responses are the main unit of analyses, see Results section), one film out of 2 was randomly selected in each repeated category for each participant. This method enabled to compute Analyses of Variance (ANOVA) in which each observation is comparable within each participant. For computation of ranking scores in film-level analyses all observations were included.

Procedure

Participants came to the laboratory in subgroups of 3 to 5 people. Upon their arrival, participants were instructed that they would see several emotional film excerpts. Participants were told that they would go through a relaxation procedure before viewing each excerpt and that they would have to report their emotional reactions to the film by filling out questionnaires after watching the excerpt. Following instructions first used by Philippot (1993), participants were encouraged to report (1) what they had actually felt and not what they believed people should feel in reaction to the movies (in order to strengthen this point,

the importance of individual differences in reactions to films was stressed) and (2) how they felt at the specific time they were watching the video excerpt, and not their general mood of the day. This procedure of retrospective evaluation of emotional feelings is known to be a good predictor of the actual state felt during the stimulation (i.e. exposure to a stimulus), especially the peak (most intense) state and the state felt by the end of the stimulation (Fredrickson & Kahnemann, 1993). Participants were also informed that their responses would remain totally anonymous. Moreover, they were told that they could stop their participation in the experiment at any time. All participants gave their consent to participate.

An individual experimental context was created for each participant: They were separated by partitions and used individual screens and headphones in order to watch the film clips without being able to see or hear the other participants' reactions. When all participants were ready to start the experiment, the room's lighting was dimmed, and pre-recorded relaxation instructions were given: Participants had to close their eyes, to relax every muscle (including the face), and breath deeply and regularly for about 2.5 min. As soon as the relaxation was over, participants were informed that the first film clip would be displayed. Participants were instructed to watch the whole scene attentively, without diverting their attention from the screen. The films were displayed on 17" color screens. After each film excerpt, participants completed questionnaires about their emotional state (see "Measures" section). This procedure (including relaxation instructions) was repeated for each film excerpt.

Measures

Self-reported emotional arousal.

The intensity of subjective emotional arousal was assessed using a 7-point scale: *"While I was watching the film… 1 = I felt no emotions at all, 7 = I felt very intense emotions."* The participants were encouraged to answer this question according to what they actually felt during the task and not according to what they thought they should feel when watching the film.

Differential Emotions Scale.

To assess discrete emotional dimensions, we used a version of The Differential Emotions Scale (DES; Izard, Dougherty, Bloxom, & Kotsch, 1974) modified by McHugo et al. (1982) and validated in French by Philippot (1993). This choice was motivated by the fact that the DES is one of the most widely used self-report scales of discrete emotional feelings (Youngstrom & Green, 2003), and this particular version of the DES has been used in two previous validations of emotional film stimuli (McHugo et al., 1982; Philippot, 1993). Each item consisted of groups of emotional adjectives: (1) *interested, concentrated, alert*; (2) *joyful, happy, amused*; (3) *sad, downhearted, blue*; (4) *angry, irritated, mad*; (5) *fearful, scared, afraid*; (6) *anxious, tense, nervous*; (7) *disgusted, turned off, repulsed*; (8) *disdainful, scornful, contemptuous*; (9) *surprised, amazed, astonished*; (10) *warmhearted, gleeful, elated*.

In order to extend the range of emotional dimensions assessed, we used a modified version of the DES in which the following items were added: (11) *loving, affectionate, friendly*; (12) *guilty, remorseful*; (13) *moved*; (14) s*atisfied, pleased*; (15) *calm*, *serene, relaxed*. (16) *Ashamed, embarrassed*. This modified version of the DES has already been successfully used in previous studies (Schaefer, et al., 2003; Schaefer & Philippot, 2005). For each group of adjectives, participants used a 7-point scale (1 = "*not at all*", 7 = "*very intense*") to rate the extent to which they felt each state as they were watching the film clip.

PANAS.

A validated French translation (Gaudreau, Sanchez & Blondin, 2006) of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) was used. The PANAS is a self-reported adjective checklist that contains two 10-item subscales designed for the assessment of Positive Affect (PA: *active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud,* and *strong*) and Negative Affect (NA: *afraid, ashamed, distressed, guilty, hostile, irritable, jittery, nervous, scared,* and *upset*). For each of the 20 emotion-related words, participants used a 5-point scale (1 = " *very slightly* or *not at all* ", 5 =

" *extremely* ") to rate the extent to which they felt each state as they were watching the film clip. Internal consistency of the PA and NA scales was assessed by Cronbach's alphas, which all indicated an acceptable level of internal consistency.

Results

Results are organised around five basic questions: (1) Which film categories yielded the strongest levels of emotional arousal? (2) Does the film set yield differentiated positive and negative affective states? (3) Can differentiated emotional feeling states be elicited with the film set?; (4) Were there any gender differences? (5) How to select efficient films from the database? Questions 1 to 4 were addressed with subject-level analyses (i.e. participants' responses are the unit of analysis across all 70 films). This allowed testing the effectiveness of the film set to elicit emotions in an experimental setting. Question 5 was addressed using film-level analyses (i.e. mean scores for each film are the unit of analysis) in order to provide a means to classify and compare individual film excerpts.

Which film categories yielded the stronger levels of emotional arousal?

A repeated-measures ANOVA tested the effect of film category (anger vs. happiness vs. fear vs. disgust vs. sadness vs. affection vs. neutral) on the arousal scale. Due to the large sample size (N=364), we decided to consider as significant only the effects that fulfilled two criteria: a *p*-value < .001, and an eta-square of at least .06 to conform to a medium effect size (Green and Salkind, 2003). A main effect of Emotion was observed, F (6, 2172) = 187.5, $p < .00001$, $\eta^2 = .34$. Post-hoc tests (only Bonferroni's at $p < .001$ were considered as significant) revealed that all comparisons involving the neutral category were highly significant at $p < .00001$ (see Fig. 1). Comparisons between emotional categories were significant, except Anger-Fear, Anger-Tenderness and Tenderness-fear. Results are depicted in Fig. 1. These results suggest that all emotional films were successful at generating stronger levels of self-reported arousal than neutral films, and they also show that

amusement films have the lowest levels of arousal (yet still higher than neutral). Further these results suggest that disgust film excerpts generated a lower arousal than the other negative film excerpts.

Insert Fig. 1 about here

Does the film set yield differentiated positive and negative affective states?

Scores were computed for Positive Affect (PA) and Negative Affect (NA) for each emotional condition, averaging the PANAS subscales' items. Two one-way repeated-measures ANOVAs were run to test the effect of film category (anger vs. happiness vs. fear vs. disgust vs. sadness vs. tenderness vs. neutral) separately on PA and NA scores across all the 70 films viewed by the subjects. We used the same significance criteria explained in the previous section. A main effect of film category was observed for both NA and PA, [NA: F (6, 2172) = 171.8, $p < .001$, $\eta^2 = .32$; PA:, F (6, 2172) = 74.6, $p < .001$, $\eta^2 = .17$]. We performed Bonferroni pairwise comparisons in order to test whether negative and positive films would be different in terms of NA and PA scores. For NA scores, negative films all had significantly higher scores than positive films ($p < 0.001$). PA scores were less successful at differentiating between positive and negative films, as Amusement-Anger, Amusement-Fear and Tenderness fear contrasts on PA scores were not significant. In order to investigate this issue, we performed the same three contrasts (Amusement-Anger, Amusement-Fear and Tenderness-Fear) for every individual item contained in the PA composite score. We found that several items were not producing unexpected results in the form of significantly higher ratings for positive than for negative films ($p < 0.001$). For the fear-tenderness comparison, this was found for "active" (M=1.8, SD = 1.0; M =1.6, SD= 0.9), "alert" (M=3.3, SD = 1.2; M =2.4, SD= 1.2) and "excited" (M=2.0, SD = 1.1; M =1.6, SD= 0.9). For the Anger-Amusement comparison, this was found for "alert" (M=3.3, SD = 1.1; M =3.1, SD= 1.1) , and for "fear-amusement", this pattern was found for "alert" (M=3.3, SD = 1.2; M =2.2, SD= 1.2), "attentive" (M=3.7, SD = 1.0; M =3.2, SD= 1.0) and "excited" (M=2.0, SD = 1.1; M =1.8, SD=

1.0). Moreover, the same contrasts were not significantly different for other sub-items of the PA scale (e.g. “determined”, “interested” and “excited”).

Taken as a whole, these results indicate that NA scores were more successful than PA scores to differentiate between positive and negative discrete emotions. The apparent failure of PA scores to differentiate negative and positive films seems to be driven by a subset of items (e.g. *attentive*, *alert*, *active*) reflecting a general “arousal” and “alertness” component which might not necessarily be valence-specific. This result argues for the need of other alternative measures to test the validity of the positive-negative differentiation of our a priori categories of emotional films. Therefore, we created two alternative negative and positive valence scores averaging classes of DES items. The positive composite score included 5 items (“*joyful, happy, amused”; “warmhearted, gleeful, elated”; “loving, affectionate, friendly”; “moved”*; “s*atisfied, pleased”*), and the negative composite score included 8 items (“*sad, downhearted, blue”; “angry, irritated, mad”*; “*fearful, scared, afraid”*; “*anxious, tense, nervous”*; “*disgusted, turned off, repulsed”*; “*disdainful, scornful, contemptuous”*; “*guilty, remorseful”*; *“Ashamed, embarrassed”*). All Cronbach’s Alphas were superior than 0.60[1.] A one-way ANOVA on the positive composite score testing for differences between emotional films revealed a significant main effect of emotional category F (5, 1810) = 289.5, $p < .001$, $\eta^2 = .44$. Bonferroni post-hoc comparisons revealed significant differences between each of the two emotional categories and all of the negative categories ($p < .001$). A one-way ANOVA on the negative composite score also revealed a main effect of emotional category F (5, 1810) = 121.1, $p < .001$, $\eta^2 = .25$, and Bonferroni post-hocs also showed a full significant differentiation between positive and negative films. These results are depicted in Figures 2-5.

Insert Figures 2, 3, 4, & 5 about here

Can differentiated emotional feeling states be elicited with the film set?

A 7X16 repeated-measures ANOVA found a significant interaction between film category and DES items, $F(90, 32670) = 228.3$, $p = 1X10^{-15}$, $\eta^2 = .39$, indicating that the self-reported emotional profile was modulated by film category. In order to further specify whether discrete emotional states were indeed induced for each film category, paired *t*-test comparisons between the DES items were computed separately for each emotional state using individual participants' self-report scores as unit of analysis. More specifically, for each category of emotional films (anger, sadness, fear, disgust, amusement, tenderness, and neutral), a set of 6 predefined paired *t*-tests contrasting the target state and each non-target state were computed for every participant. For instance: for the category of films evoking *fear*, one target DES item was chosen (*fearful, scared, afraid*) and compared to 6 non-target items, which are targets to the other 6 emotional categories (namely: anger, sadness, amusement, tenderness, neutral, disgust). Due to the large sample size (N=364) and to the large number of planned comparisons, we decided to (a) apply Bonferroni corrections for multiple comparisons, and (b) to set a stringent significance threshold of $p < 0.0001$. Therefore, only *t*-values > 4.3 were considered significant. Results revealed that most comparisons were significant according to these criteria, confirming the expected differentiation between target states. Only one exception was found: The difference between fear and anger did not reach the significance criterion during anger induction (although the same difference is highly significant during fear induction). To further analyze this issue, we computed discreteness analyses separately for every anger, film. The five film excerpts with the higher level of anger discreteness (see the "Ranking Scores" section for the calculation of this coefficient), yielded significant anger-fear comparisons ($p<0.005$) for all but one of them (*American History X*), indicating that a fairly good number of anger excerpts yielded a high level of differentiation between anger and fear when they were analyzed individually. Individual film identities can be found in Table 1.

Were there any gender differences?

In order to assess gender differences, ANOVAs described above were run with gender as an additional factor. We found a main effect of Gender in the arousal scale, F (1, 362) = 30.1, $p < .0001$, $\eta^2 = .08$, suggesting that women reported higher levels of emotional arousal than men when viewing video-clips (Women: $M = 4.4$, $SE = 0.04$, Men: $M = 3.8$, $SE = 0.09$). The Emotion X Gender interaction failed to pass the significance criteria. For NA, a main effect of Gender was found, F (1, 362) = 28.7, $p < .001$, $\eta^2 = .07$, indicating that women reported a higher NA level than men when viewing video clips (Women: $M = 1.7$, $SE = 0.02$, Men: $M = 1.4$, $SE = 0.04$). The Emotion X Gender interaction failed to pass the significance criteria. No effects of Gender passed the significance criteria in analyses involving the PA factor or DES measures.

How to select efficient films from the database?

A film-level analysis was computed in order to provide a means to classify and compare individual films according to specific criteria. We therefore computed for each film a mean of each measured variable, obtaining a final data matrix including the scores of each film for 19 variables corresponding to the arousal scale, the positive and negative affect scores, and the 16 extended-DES items. Next, we computed two film-level coefficients for positive and negative composite scores averaging classes of DES items as explained previously (Question 2 of the Results section). Finally, we computed two classes of coefficients aimed to classify film stimuli according the levels of differentiation of emotional states elicited by them. First, we computed discreteness coefficients for each emotional category: the mean DES score of the scale targeting one particular emotion minus the averaged mean scores of the scales targeting the other five emotions. Hence, we obtained six discreteness coefficients, one for each emotion (amusement, tenderness, anger, sadness, disgust and fear). Next, we also computed a “mixed feelings” (MF) coefficient which estimates how mixed/blended were the feelings elicited by each film[2]. We used a formula used in previous research (Hemenover & Schimmack, 2007; Schimmack, 2001), in which the

MF score is computed as the minimum of two affect scores elicited by the same stimulus. For instance, if film Y received a score of 5 in the “Anger” item and a score of 2 in the “Fear” item, then MF = 2. More positive values are associated with a higher level of mixed Anger-Fear reported states. The calculation of MF scores resulted in 15 additional criteria corresponding to the pairwise combinations of all the emotional states investigated in this study (Anger, Sadness, Fear, Disgust, Amusement, Tenderness). It is important to note that discreteness and mixed feelings scores are measuring two different constructs. The former are measuring how much a specific state is activated relative to an average of all other states. The latter are measuring how much two specific states are mixed/differentiated on a pairwise basis.

The resulting final 26 ranking criteria were then: (1) The subjective arousal score, (2) PA and NA scores, (3) a positive and a negative composite scores (DES-based), (4) the six discreteness scores, (5) the 15 mixed feelings scores. We defined subgroups of high-ranking films for the 26 final criteria in order to identify individual films with high levels of efficiency (see “High-scoring subgroups” section). Within each subgroup, an outlier analysis was performed in order to detect whether specific films were less or more efficient than the other films of the subgroup. Films included in high-scoring subgroups for Arousal, PA, NA, composite scores and discreteness scores are listed in Table 1. For conciseness, we report in detail the mixed feeling scores in our companion website (http://nemo.psp.ucl.ac.be/FilmStim/). We describe in the next section how the effectiveness of high-ranking films was ascertained.

Finally, we selected four neutral films from the original database that could serve as control films in most emotion studies. These film clips had the lowest arousal, PA and NA scores in the entire database, and they are brief excerpts of international films easily available in several languages (neutral clips taken from local Belgian TV programs were not included in the final set of recommended film clips). The four recommended neutral clips are the following: “Blue (1)”, “Blue (2)”, “Blue (3)” and “The lover”..

High-scoring subgroups.

We defined subgroups of particularly efficient films for each criterion. We first identified the 10 highest-ranking films for each criterion. Next, for each criterion, we tested whether each high-ranking film was different from the average of the entire database of films. All comparisons were significant (all *p*s < .0001) for all criteria, indicating that every film included in the subgroups presented in Table 1 has a high level of efficiency for one of the 26 validity criteria. As a further check, we proceeded to an outlier analysis within each high-scoring subgroup in order to detect specific films that would have a significantly different level of efficiency than the other films of the subgroup. This analysis was an additional means to filter out any potentially inefficient film. Outliers were defined as scores beyond 1.5 interquartile ranges from the 1st and 3rd quartile of each subgroup (Tukey, 1977). No subgroup was found to include films significantly below the subgroup median score, and the distribution within subgroups appeared to be homogeneous (Figure 6). Observed outliers were for films obtaining scores well above the median of their subgroups, indicating that their efficiency is even higher than the efficiency of the high-scoring films. These films were the following: “The dead Poets Society (2) “, “Trainspotting (2)” and “American History X” were found to be outliers for PA, NA and disgust discreteness, respectively. For mixed-feelings scores, “Life is Beautiful (4) was an outlier for Amusement-Tenderness, “Man bites dog (2)” was an outlier for Amusement-Anger, “Trainspotting (3)” was an outlier for Amusement-Disgust, “A perfect world” and “Life is beautiful (1)” were outliers for both Tenderness-Anger and Tenderness-Disgust, and “Schindler’s list (1)” was an outlier for Sadness-Fear.

Insert Fig. 6 and Table 1 about here

Discussion

The present study tested the effectiveness of a set of 70 film clips to elicit emotions in an experimental context according to different validity criteria. Several questions were addressed by the data analysis. First, ANOVAs showed that the films were capable of eliciting significantly intense emotions and differentiated responses in terms of positive and

negative affect. Second, Bonferroni comparisons yielded fairly high levels of discreteness for the 7 target states for each category of film.

Several validity criteria were defined to classify and compare films: A general arousal criterion, seven criteria for emotional discreteness (Anger, sadness, fear, disgust, amusement, tenderness), a general positive affect dimension (PANAS positive affect scale), a general negative affect dimension (PANAS negative affect), two positive and negative composite scores based on the DES, and 15 "mixed feelings" scores. For each criterion, a subgroup was created. Results indicated that films contained in these subgroups are the most likely to efficiently elicit expected affective states in laboratory sessions. These films are presented in Table 1, and extensive additional data on these films and on all the films of the database are available on Internet[3] along with descriptions of the film excerpts and editing instructions. This allows emotion researchers to select the most appropriate stimuli for their research and to edit the excerpts in versions of the movies dubbed in their study participants' language.

A notable unexpected result was the failure of the PA scale (Watson & Tellegen, 1988) to differentiate negative and positive films. This phenomenon was already observed with emotional picture stimuli (Patrick & Lavoro, 1997) in a study showing that negative pictures were generating high PA scores. An analysis of individual items showed that this effect was driven in our data by specific items which appear to reflect a valence-independent arousal/alertness component likely to be highly activated in fear, anxiety and anger states. Interestingly, the same items were also driving Patrick & Lavoro's (1997) results. Therefore, caution has to be recommended when using the PANAS scales in order to measure participants' responses to emotional stimuli. We have calculated alternative positive and negative valence scores (based on the DES), which were successful in differentiating positive and negative films. These alternative scales were also provided as classification criteria.

Next, gender differences in reactivity to the films were observed. More specifically, women seemed to report higher levels of subjective arousal, and higher levels of negative

affect than men. No significant gender differences were found in positive affect as defined by the PANAS scales. A higher level of reported emotional feelings for female participants is consistent with Gross & Levenson (1995) and Hagemann et al. (1999) results and coherent with many studies showing that women usually report more intense emotions than men (Brody & Hall, 2000). These results may reflect either neurobiological differences in emotional systems across genders (Cahill et al., 2004) or the effects of socio-cultural stereotypes about gender roles (Brody & Hall, 2000; Feldman-Barrett et al., 1998).

Two caveats concerning this study have to be considered. First, this validation study was conducted in a specific cultural context (French-speaking European) that may have influenced the choice and evaluation of films for some emotional categories. In order to address this issue, we included films from different cultural backgrounds (French, Italian, British and US cinema) to allow some culture-specific choice of stimuli. For instance, even though some amusement films are derived from the French cinema, half of the ten highest ranked amusement excerpts, which were taken from US and British cinema: *"When Harry met Sally"* , *"A fish called Wanda", "Benny and Joon" and "Something about Mary"*. Further research will be needed to expand the choice of stimuli to non-western cultures. Second, in line with all previous film validation experiments, the present study relied solely on self-reports to assess emotional responses. Future research is needed to validate stimuli on expressive and psychophysiological criteria.

Beyond using films as elicitors of affective states that can modulate cognition and social behavior, we believe that several pathways are open for future research on emotional responsiveness to films *per se*. For instance, in combination with psychophysiological and neuroimaging techniques (e.g. fMRI, ERP, MEG, eye-tracking techniques, etc.) research using film stimuli may yield interesting insights about the perceptual and attentional processes involved in emotional processing, as well as the neuroanatomical and neurochemical correlates of emotional processes. Likewise, using film excerpts could help further our understanding of the time course of an emotional state, i.e. how an emotional reaction unfolds over a given period, and what factors determine its maintenance or decline.

This question is linked to the duration of an emotional state, a feature important for conceptual distinctions between mood and emotion (Ekman & Davidson, 1994, pp. 49-96 ; Frijda, 1993).

This study attempted to validate emotional film stimuli taking in account the theoretical diversity of the field. Specifically, we defined several validity criteria derived from two categories of theoretical models of emotion: the basic emotions approach and the dimensional approach. Although these two theories are prominent, they do not exhaust the diversity of theories of emotion. Of a particular interest is the componential approach – or appraisal models, that define emotion in terms of different combinations and blends of discrete cognitive and motivational components such as goal-relevance, novelty and expectancy (Scherer, 1999). Future developments in emotion elicitation techniques would potentially benefit from using validity criteria derived from these constructs.

In sum, the main goal of this study was to build a *flexible* tool for emotion research, that is, a set of film clips from which different subsets of stimuli can be chosen according to specific criteria. For this purpose, we gathered a comprehensive set of films which successfully fulfilled different validity criteria. Next, we created several different indices along which films were ranked. The entire database was made freely available on Internet, including the film-level data matrix so that anyone can download and analyze the data in order to flexibly select a set of stimuli. It is worth noting that data from additional film excerpts can be added by researchers who wish to validate other film excerpts using similar methods and measures. The comparison with the original database could provide an index of external validity to the newly added film excerpts.

References

Averill, J.R. (1982). On the paucity of positive emotions. In K.R. Blankstein, P. Pliner, & J. Polivy (Eds.), *Assessment and modification of emotional behavior*. New York, Hillsdale: Plenum Press.

Bradley, M.M. & Lang, P.J. (2000). Measuring emotion: behavior, feeling and physiology. In R.D. Lane and L.Nadel (Eds.), *Cognitive Neuroscience of emotion*. (pp. 242-276). Oxford University Press.

Brody, L.R. & Hall, J.A. (2000). Gender, emotion and expression. In M. Lewis and J.M. Haviland-Jones (Eds.). *Handbook of emotions*. Guilford Press.

Cahill, L. Uncapher, M., Kilpatrick, C., Alkire, M.T., Turner, J. (2004). Sex-Related Hemispheric Lateralization of Amygdala Function in Emotionally Influenced Memory: An fMRI Investigation. Learning & Memory, 11:261–266.

Davidson, R.J. (1992). A prolegomenon to the structure of emotion: gleanings from neuropsychology. *Cognition and Emotion*, 6, 245-268.

Davidson, R.J. (1993). Parsing affective space: Perspectives from neuropsychology and psychophysiology. Special section: Neuropsychological perspectives on components of emotional processing. *Neuropsychology, 7*, 464-475.

Davies, C. (1998). The dog that didn't bark in the night: A new sociological approach to the cross-cultural study of humor. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (Humor Research Series, vol. 3, pp. 293-306). Berlin: Mouton de Gruyter.

Depue, R.A. & Morrine-Strupinsky, J.V. (In Press). A neurobehavioral model of affiliative bonding: implications for conceptualizing a human trait of affiliation. Behavioral and Brain Sciences.

Ekman, P. (1984). Expression and the nature of emotion. In K.R. Scherer & P. Ekman (Eds.) *Approaches to emotion* (pp. 319-344). Hillsdale, NJ: Erlbaum.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169-200.

Ekman, P. & Davidson, R.J. (1994). *The nature of emotion: fundamental questions.* Oxford University Press.

Feldman-Barrett, L., Robin, L, Pietromonaco, P.R. & Eyssel, K. M. (1998). Are women the more emotional sex? Evidence from emotional experiences in social context. Cognition and Emotion, 12, 555-578.

Fredrickson, B. L. and Kahneman, D., 1993. Duration neglect in retrospective evaluations of affective episodes. J Pers Soc Psychol. 65, 45-55.

Frazier, T.W. ; Strauss, M.E. ; Steinhauer, S.R. (2004). Respiratory sinus arrhythmia as an index of emotional response in young adults. *Psychophysiology* 41(1): 75-83.

Frijda, N.H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Frijda, N. H. (1993). Mood, emotion episodes, and emotions. In M. Lewis & J. M. Haviland (Eds) Handbook of emotions (pp. 381-404). New York: The Guilford Press.

Gaudreau, P., Sanchez, X. & Blondin, J.P. (2006). Positive and negative affective states in a performance-related setting: Testing the factorial validity of the PANAS across two samples of French-Canadian participants. *European Journal of Psychological Assessment*, 22, 240-249.

Green, S. B., & Salkind, N. J. (2003). *Using SPSS: Analyzing and understanding data* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Gross, J.J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. Journal of Personality and Social Psychology, 74, 224-237.

Gross, J.J., & Levenson, R.W. (1995). Emotion elicitation using films. Cognition and Emotion, 9, 87-108.

Hagemann, D.; Naumann, E.; Maier, S.; Becker, G.; Lurken, A.; Bartussek, D. (1999). The assessment of affective reactivity using films: validity, reliability and sex differences. *Personality and Individual Differences*, 26, 627 – 639.

Hatfield, E. & Rapson, (2000). R.L. Love and attachment processes. In M. Lewis and J.M. Haviland-Jones (Eds.). *Handbook of emotions*. Guilford Press.

Hemenover, S. H. & Schimmack, U. (2007). That's disgusting! …, but very amusing: Mixed feelings of amusement and disgust. *Cognition and Emotion*, 21 (5), 1102-1113.

Izard, C.E. (1991). The psychology of emotions. New York: Plenum Press.

Izard, C.E., Dougherty, F.E., Bloxom, B.M., & Kotsch, N.E. (1974). *The Differential Emotions Scale: A method of measuring the meaning of subjective experience of discrete emotions.* Nashville: Vanderbilt University, Department of Psychology.

Kelly, J.R. & Hutson-Corneaux, S.L. (1999). Gender emotion stereotypes are context specific. *Sex Roles, 40*, 107-120.

Landis, C. (1924). Studies of emotional reactions. II. General behavior and facial expression. *Journal of Comparative Psychology*, 4, 447-509.

Lang, P.J., Greenwald, M.K., Bradley, M.M. & Hamm, A.O. (1993). Looking at pictures: Affective, facial, visceral and behavioral reactions. *Psychophysiology*, 30, 261-273.

Lazarus, R.S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist, 46*, 819-834.

Lazarus, R.S. and Smith, C.A. (1991). Emotion and adaptation. In L.A. Pervin (Ed.). *Handbook of personality: theory and research*. (pp. 609-637). NY: Guilford Press.

Le Doux, J.E. (1996). *The emotional brain: the mysterious underpinnings of emotional life.* New York: Simon & Schuster.

Mathews, A.; Mackintosh, B..and Fulcher, E.P. (1997) Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences*, 1, 340-345.

Matsumoto, D. (1987). The role of facial responses in the experience of emotion: More methodological problems and a meta-analysis. *Journal of Personality and Social Psychology*, 52, 769-774.

McHugo, G.J. ; Smith, C.A. ; Lanzetta, J.T. (1982). The structure of self-reports of emotional responses to film segments. *Motivation & Emotion* 6(4): 365-385.

Palomba, D. ; Sarlo, M. ; Angrilli, A. ; Mini, A. ; Stegagno, L. (2000). Cardiac responses associated with affective processing of unpleasant film stimuli. *International Journal of Psychophysiology* 36(1): 45-57.

Panksepp, J. (1998). *Affective Neuroscience*. New York: Oxford University Press.

Patrick, C.J. & Lavoro, S.A. (1997). Ratings of emotional response to pictorial stimuli: positive and negative affect dimensions. *Motivation and Emotion*, 21(4), 297-321.

Philippot, P. (1993). Inducing and assessing differentiated emotional feeling states in the laboratory. Cognition and Emotion, 7, 171-193.

Philippot, P., Chapelle, C. & Blairy, S. (2002). Respiratory feedback in the generation of emotion. *Cognition & Emotion*, 16, 605-627.

Philippot, P., Schaefer, A., & Herbette, G. (2003). Consequences of specific processing of emotional information: Impact of general versus specific autobiographical memory priming on emotion elicitation. *Emotion*, 3, 270-283.

Plutchik, R. (1980). Emotion: A psychobioevolutionary synthesis. New York: Harper & Row.

Rottenstreich, Y. & Hsee, C. K. (2001). Money, kisses and electric shocks: On the affective psychology of risk. *Psychological Science*, 12, 185-190.

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

Schaefer, A.; Braver, T.S.; Reynolds, J. R.; Burgess, G.C.; Yarkoni, T.; & Gray, J.R. (2006). Individual differences in amygdala activity predict response speed during working memory. *Journal of Neuroscience*, 26(40): 10120-10128.

Schaefer A.; Fletcher, K.; Pottage, C.; Alexander, K. & Brown, C. (2009). The effects of emotional intensity on ERP correlates of recognition memory. *Neuroreport, 20(3).*

Schaefer, A. & Philippot, P. (2005). Selective effects of emotion on the phenomenal characteristics of autobiographical memories. *Memory, 13,* 148-161.

Schaefer, A., Collette, F., Philippot, P., Vanderlinden, M., Laureys, S., Delfiore, G., Degueldre, C., Maquet, P., Luxen, A. & Salmon, E. (2003). Neural correlates of "hot"

and "cold" emotional processing: A multilevel approach to the functional anatomy of emotions. *Neuroimage, 18*, 938-949.

Scherer, K.R. (1999). Appraisal theory. In T. Dalgleish and M. Power (Eds.), *Handbook of cognition and emotion*. (pp. 637-663).UK: Wiley.

Scherer, K.R. (2000). Psychological models of emotion. In J.C. Borod (Ed.), *The neuropsychology of emotion*. (pp. 137-162).Oxford University Press.

Schimmack, U. (2001). Pleasure, displeasure, and mixed feelings: Are semantic opposites mutually exclusive? *Cognition and Emotion*, *15* (1), 81–97.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*, 350-353.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: further exploration of a prototype approach. *Journal of Personality and Social Psychology, 52*, 1061-1086.

Smith, C.S. & Lazarus, R.S. (1993). Appraisal components, core relational themes, and the emotions. *Cognition and Emotion*, *7*, 233-269.

Sonnemans, J.; Frijda, N. H. (1995). The determinants of subjective emotional intensity. Cognition & Emotion 9(5): 483-506.

Stemmler, G., Heldmann, M., Pauls, C.A., & Scherer, T. (2001). Constraints for emotion specificity in fear and anger: The context counts. Psychophysiology, 38, 275-291.

Tamaoka, K. & Takahashi, T. (1994). Understanding humour from another culture: Comprehension of parental brain twisters by Japanese university students learning English as a second language. *Psychologia* 37(3): 150-157.

Tukey, J.W. (1977). Explanatory data analysis. Reading, MA: Addison-Wesley.

Velten, E. (1968). A laboratory task for induction of mood states. *Behavior Research and Therapy*, 6, 473-482.

Vrana, S.R., Cuthbert, B.N., Lang, P.J. (1986). Fear imagery and text processing. Psychophysiology, 23, 247-253.

Watson, D. ; Clark, L. A. ; Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality & Social Psychology* 54(6): 1063-1070.

Westermann, R.; Spies, K.; Stahl, G. & Hesse (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology,* 26(4), 557-580.

Youngstrom, E.A. & Green, K.W. (2003). Reliability Generalization Of Self-Report Of Emotions When Using The Differential Emotions Scale. *Educational and Psychological Measurement* 2003; 63; 279-295.

Footnotes

1. We used the criterion proposed by Schmitt (1996), according to which an α = .50 is satisfactory if the composite measure has a meaningful content coverage of some domain and reasonable unidimensionality.

2. We thank Scott Hemenover for this suggestion.

3. Detailed information about the film excerpts used in this study is available at the following URL: http://nemo.psp.ucl.ac.be/FilmStim/. This website contains the database used for this study, free of charge.

Table 1: Table 1. Sets of ten first ranked film excerpts.

Rank	Arousal	Positive Affect	Negative Affect
1	*Misery* (6.12) Annie (Kathy Bates) breaks Paul's legs (James Caan).	*The Dead Poets Society [2]* (2.82) All the students climb on their desks to manifest their solidarity with Mr. Keating (Robin William), who has just been fired.	*American History X* (2.73) A neo-nazi (played by Edward Norton) kills an afro-american man smashing his head on the curb.
2	*Schindler's List [1]* (5.88) Dead bodies are being carried away in a concentration camp.	*The Dinner Game* (2.50) Complex humouristic scene.	*Schindler's List [1]* (2.62) Dead bodies are being carried away in a concentration camp.
3	*Dead Man Walking* (5.87) The main character is put to death by lethal injection.	*Life is beautiful [4]* (2.49) In a concentration camp, a father "fakes" a translation of what an officer says in order to prevent his son to be frightened.	*The Piano* (2.49) One of the characters gets her hand cut off.
4	*American History X* (5.84) A neo-nazi (played by Edward Norton) kills an afro-american man smashing his head on the curb.	*Forrest Gump* (2.37) Father and son are reunited.	*Misery* (2.49) Annie (Kathy Bates) breaks Paul's legs (James Caan).
5	*Life is Beautiful [2]* (5.82) In a prisoner's camp, a father (Roberto Benigni) and a boy talk to the mother using a loud speaker, reaching the whole camp.	*Life is Beautiful [2]* (2.34) In a prisoner's camp, a father (Roberto Benigni) and a boy talk to the mother using a loud speaker, reaching the whole camp.	*Sleepers* (2.46) Sexual abuse of children.
6	*A perfect World* (5.78) Butch (Kevin Costner) is gunned down, at the end of the movie.	*When Harry met Sally* (2.32) Sally simulates an orgasm in a restaurant.	*Scream 1* (2.43) A girl receives threats through the phone
7	*Saving Private Ryan* (5.73) Graphic war scene: fighting on the beaches.	*"Les trois frères"* (2.32) One of the characters takes part in a TV game.	*Schindler's list [2]* (2.43) A concentration camp commander randomly shoots prisoners from his balcony.
8	*Seven [1]* (5.69) By the end of the movie, Kevin Spacey tells Brad Pitt that he beheaded his pregnant wife.	*Benny and Joone* (2.32) Benny (Johnny Depp) plays the fool in a coffee shop.	*The Blair Witch Project* (2.40) Final scene in which the characters are apparently killed.
9	*The Piano* (5.67) One of the characters gets her hand cut off.	*"La cité de la peur"* (2.31) Conversation between three characters at a dinner table.	*The Shining* (2.35) The character played by Jack Nicholson pursues his wife with an axe.
10	*The Dead Poets Society [2]* (5.66) All the students climb on their desks to manifest their solidarity with Mr. Keating (Robin William), who has just been fired.	*Life is Beautiful [3]* (2.28) Mother and son are reunited.	*Saving Private Ryan* (2.33) Graphic war scene: fighting on the beaches.

Table 1. Sets of ten first ranked film excerpts. (continuation).

Rank	Amusement (*d*)	Tenderness (*d*)	Anger (*d*)
1	*"Les trois frères"* (4.55) One of the characters takes part in a TV game.	*Forrest Gump* (4.20) Father and son are reunited.	*Schindler's List [2]* (2.19) A concentration camp commander randomly shoots prisoners from his balcony.
2	*The Dinner Game* (4.50) Complex humouristic scene.	*Life is Beautiful [2]* (4.17) In a prisoner's camp, a father (Roberto Benigni) and a boy talk to the mother using a loud speaker, reaching the whole camp.	*Sleepers* (2.14) Sexual abuse of children.
3	*'' La cité de la peur''* (4.29) Converstaion between three characters at a dinner table.	*Life is Beautiful [3]* (3.87) Mother and son are reunited.	*Leaving Las Vegas* (2.00) The main character is raped and beaten by three drunk men.
4	*The Visitors* (4.27) Two men wearing medieval armours attack the postman's car.	*The Eight Day* (3.61) Scene of friendship between the two main characters.	*American History X* (1.94) A neo-nazi (played by Edward Norton) kills an afro-american man smashing his head on the curb.
5	*A Fish Called Wanda* (3.81) One of the characters (John Cleese) is found naked by the owners of the house.	*Life is Beautiful [4]* (3.60) In a concentration camp, a father "fakes" a translation of what an officer says in order to prevent his son to be frightened.	*Schindler's List [1]* (1.70) Dead bodies are being carried away in a concentration camp.
6	*Benny and Joone* (3.73) Benny (Johnny Depp) plays the fool in a coffee shop.	*The Dead Poets Society [2]* (3.24) By the end of the movie, all the students climb on their desks to manifest their solidarity with Mr. Keating (Robin William), who has just been fired.	*Schindler's List [3]* (1.33) Killing of Jews in a ghetto during WWII.
7	*Something about Mary [2]* (3.68) Mary (Cameron Diaz) takes sperm from Ted's hair (Ben Stiller) mistaking it for hair gel.	*Ghost* (3.00) The ''pottery'' scene.	*Man Bites Dog [1]* (1.30) Collective rape scene.
8	*Something about Mary [1]* (3.61) Ben Stiller fights with a dog.	*E.T.* (2.81) E.T. is apparently dying.	*In the name of the Father* (1.10) Violent police interrogation leading to forged confessions.
9	*When Harry met Sally* (3.55) Sally simulates an orgasm in a restaurant.	*When a man loves a woman* (2.79) Reconciliation between two lovers.	*Seven [1]* (0.99) By the end of the movie, Kevin Spacey tells Brad Pitt that he beheaded his pregnant wife.
10	*"Le Pari"* (3.48) Lunch to celebrate father's birthday.	*The Professional [1]* (2.72) The two main characters are separated forever.	*A perfect world* (0.96) Butch (Kevin Costner) is gunned down, at the end of the movie.

Table 1. Sets of ten first ranked film excerpts. (continuation).

Rank	Sadness (*d*)	Fear (*d*)	Disgust (*d*)
1	*City of Angels* (2.32) Maggie (Meg Ryan) dies in Seth's (Nicolas Cage) arms.	*The Blair Witch Project* (2.93) Final scene in which the characters are apparently killed.	*Trainspotting [2]* (4.07) The main character dives into a filthy toilet.
2	*Dangerous Mind* (2.29) Students in a school class are told that one of their classmates has died.	*The Shining* (2.77) The character played by Jack Nicholson pursues his wife with an axe.	*Seven [3]* (3.41) Policemen find the body of a man tied to a table.
3	*Philadelphia* (2.27) Andrew (Tom Hanks) and Joe (Denzel Washington) listen to an opera aria on the stereo. Ted describes to Joe the pain and passion felt by the opera character.	*Scream 1* (2.38) A girl receives threats through the phone.	*Hellraiser* (3.14) On the floor, the size of two stains are growing, and progressively transforming into a monster with a human-like skeleton.
4	*The Dead Poets Society [1]* (1.64) A schoolboy commits suicide.	*Misery* (2.32) Annie (Kathy Bates) breaks Paul's legs (James Caan).	*Man Bites Dog [1]* (2.99) Collective rape scene.
5	*The Dreamlife of Angels* (1.46) One of the characters commits suicide.	*Scream 2* (2.12) A pursuit takes place in a school.	*Seven [2]* (2.90) Policemen find the body of a savagely tortured man.
6	*Dead Man Walking* (1.02) The main character is put to death by lethal injection.	*'Child's Play 2: Chucky's Back'* (2.02) Chucky beats Andy's teacher with a ruler.	*The Dentist* (2.84) A man finds a woman whose tongue has been savagely cut off.
7	*The Professional [1]* (0.93) The two main characters are separated forever.	*Copycat* (1.96) One of the characters gets caught by a murderer in a toilet.	*American History X* (2.77) A neo-nazi (played by Edward Norton) kills an afro-american man smashing his head on the curb.
8	*Schindler's list [1]* (0.88) Dead bodies are being carried away in a concentration camp.	*The Dentist* (1.91) A man finds a woman whose tongue has been savagely cut off.	*Sleepers* (2.76) Sexual abuse of children.
9	*A Perfect World* (0.84) Butch (Kevin Costner) is gunned down, at the end of the movie.	*The Exorcist* (1.83) A priest tries to cure a girl who is apparently possessed by the devil.	*Misery* (2.69) Annie (Kathy Bates) breaks Paul's legs (James Caan).
10	*Life is Beautiful [1]* (0.81) The main character is killed.	*Seven [2]* (1.81) Policemen find the body of a savagely tortured man.	*The Silence of the Lambs* (2.54) Forensic examination of a dead body.

Table 1. Sets of ten first ranked film excerpts. (continuation).

Rank	Positive score	Negative score
1	*Life is Beautiful [4]* (4.18)	*Schindler's List [1]* (4.11)
	In a concentration camp, a father "fakes" a translation of what an officer says in order to prevent his son to be frightened.	Dead bodies are being carried away in a concentration camp.
2	*Forrest Gump* (3.96)	*American History X* (3.84)
	Father and son are reunited.	A neo-nazi (played by Edward Norton) kills an afro-american man smashing his head on the curb.
3	*The Dinner Game* (3.89)	*Sleepers* (3.70)
	Complex humouristic scene.	Sexual abuse of children.
4	*Ghost* (3.88)	*Schindler's List [2]* (3.61)
	The "pottery" scene.	A concentration camp commander randomly shoots prisoners from his balcony.
5	*Life is Beautiful [2]* (3.87)	*Dead Man Walking* (3.54)
	In a prisoner's camp, a father (Roberto Benigni) and a boy talk to the mother using a loud speaker, reaching the whole camp.	The main character is put to death by lethal injection.
6	*The Dead Poets Society [2]* (3.86)	*Man Bites Dog [1]* (3.53)
	By the end of the movie, all the students climb on their desks to manifest their solidarity with Mr. Keating (Robin William), who has just been fired.	Collective rape scene.
7	*Benny and Joone* (3.80)	*The Piano* (3.35)
	Benny (Johnny Depp) plays the fool in a coffee shop.	One of the characters gets her hand cut off.
8	*Life is Beautiful [3]* (3.75)	*Saving Private Ryan* (3.31)
	Mother and son are reunited.	Graphic war scene: fighting on the beaches.
9	*"Les trois freres"* (3.69)	*Leaving Las Vegas* (3.29)
	One of the characters takes part in a TV game.	The main character is raped and beaten by three drunk men.
10	*The Eigth Day* (3.60)	*Misery* (3.21)
	Scene of friendship between the two main characters.	Annie (Kathy Bates) breaks Paul's legs (James Caan).

Notes. In each column, the films listed correspond to the 10 movies that obtained the highest score on a given criterion. The ranks are in a descending order (e.g. *Misery* is the most intense film excerpt) with mean scores in parentheses. *d* indicates that movies are ranked according to their score of discreteness in a given emotional dimension. Numbers in square brackets correspond to different scenes of the same film[3]. Mixed feelings coefficients are available online http://nemo.psp.ucl.ac.be/FilmStim/

Figure captions

Figure 1. Subject-level mean self-reported emotional arousal as a function of emotional category. Error bars represent ± 1 standard deviation of the mean. AM=Amusement; TEND=Tenderness; DISG=Disgust, ANG=Anger; SAD=Sadness; NEUT=Neutral.

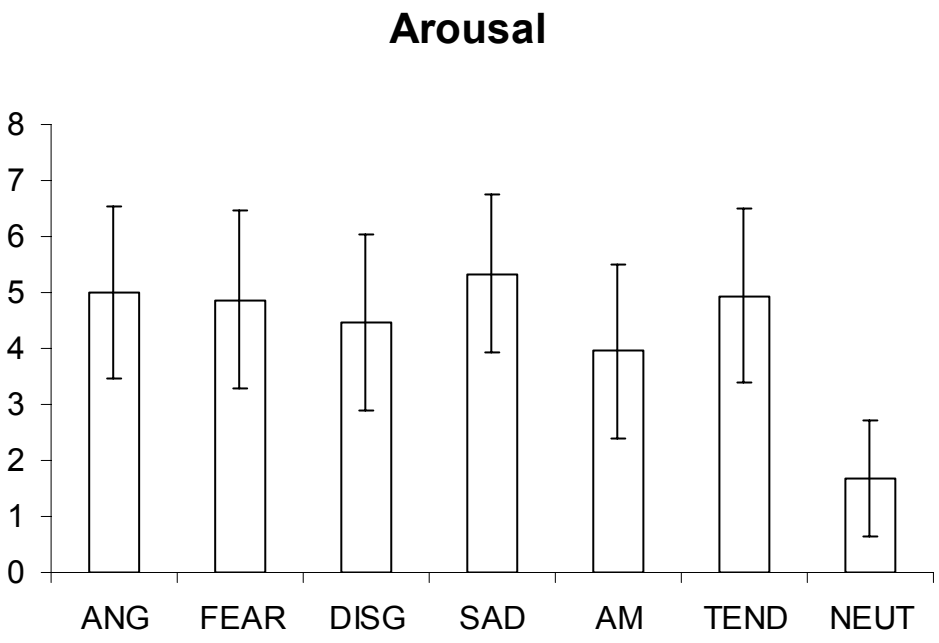
Figure 2. Positive Affect (Watson et al., 1988): Subject-level means and standard deviations

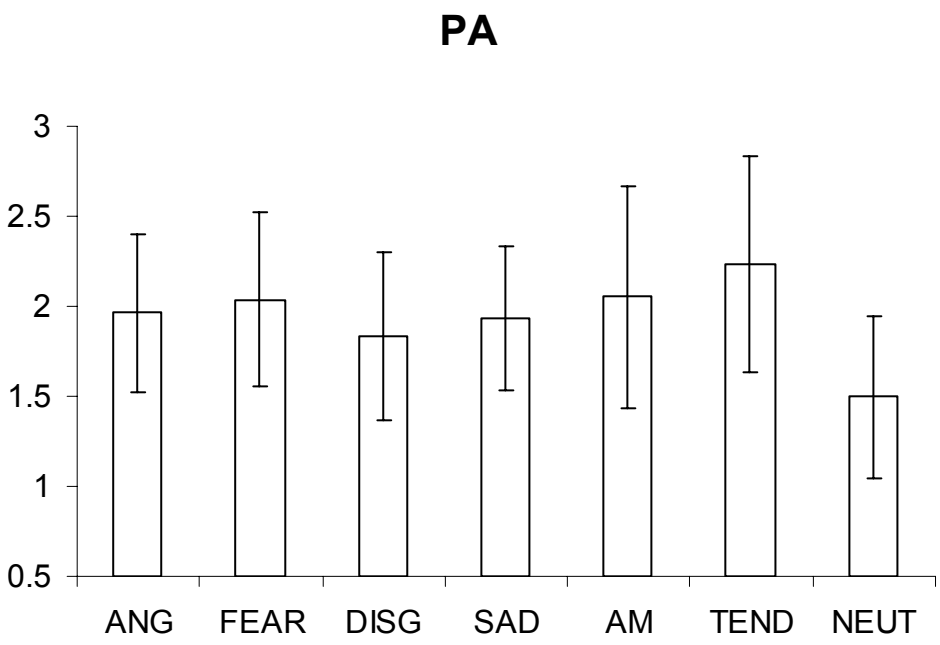
Figure 3. Negative Affect (Watson et al., 1988) : Subject-level means and standard deviations

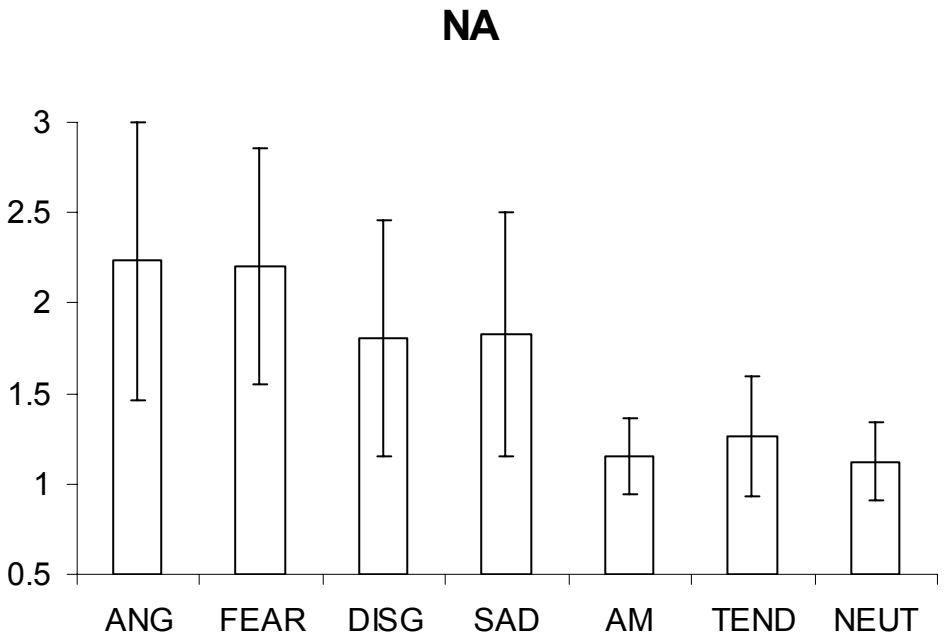
Figure 4. Positive composite score (DES-based) : Subject-level means and standard deviations

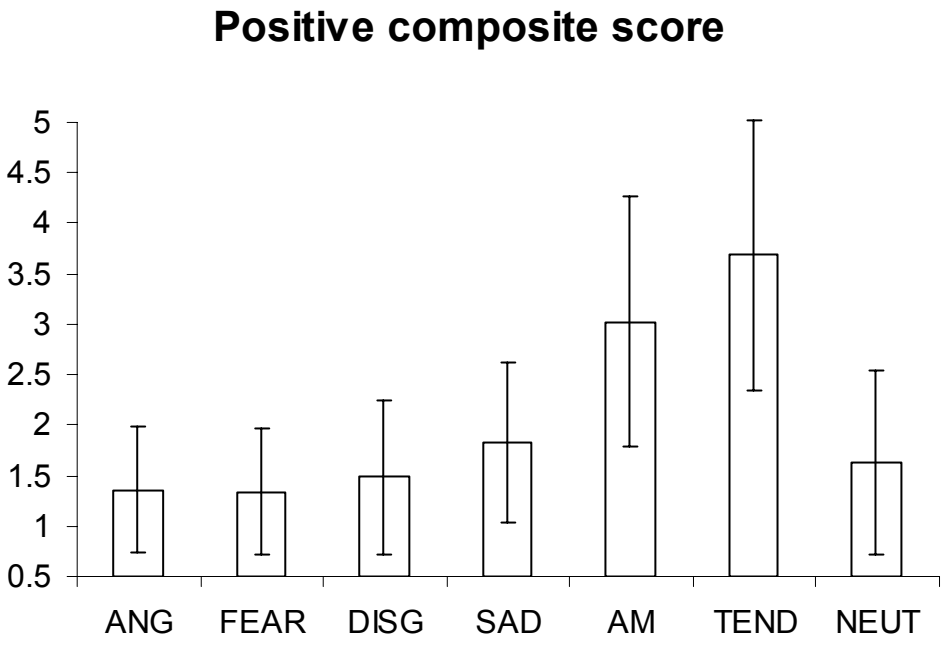
Figure 5. Negative composite score (DES-based) : Subject-level means and standard deviations

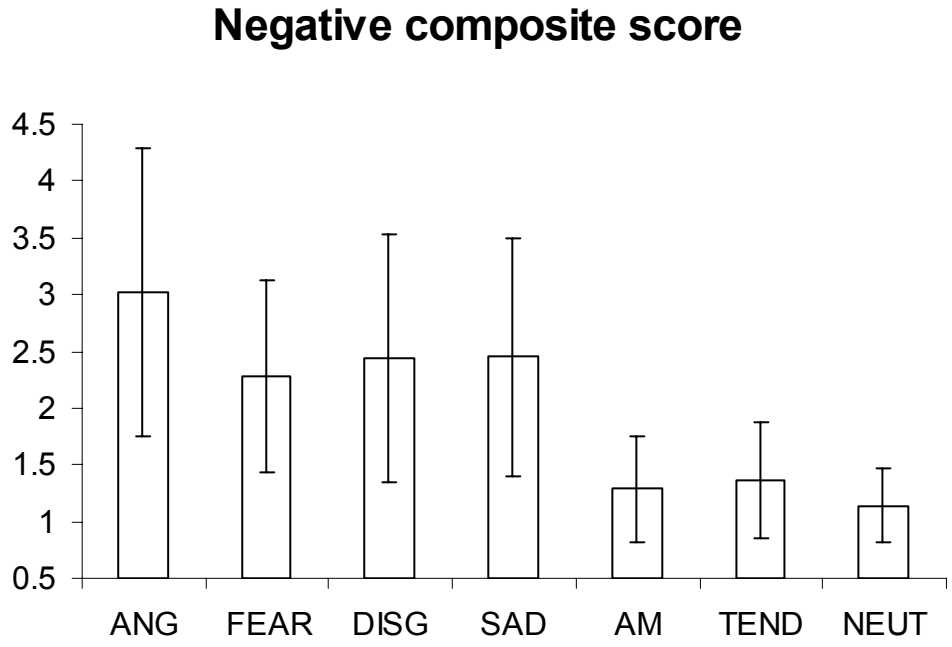
Figure 6. Distribution boxplots of subgroups of high-scoring films and outliers for each validity criterion. The vertical axis represents a scale of standardized scores (Z-scores) for all the classification criteria. Each grey box represents the interquartile range of the subgroup; the dark horizontal line represents the median of the subgroup; the whiskers indicate the farthest scores which are not outliers and the dots designate outliers. An outlier is a score at 1.5 interquartile ranges from the lower and upper edges of the box. A star represents scores beyond 3 interquartile ranges. The identity of the outliers is listed in the text. Numbers in the horizontal axis refer to specific classification criteria: 1 = subjective arousal; 2 = NA: PANAS negative affect; 3 = PA: PANAS positive affect; 4 = Positive composite score (DES-based); 5 = Negative composite score. The following numbers refer to discreteness coefficients: 6 =Amusement; 7 = Tenderness; 8 = Anger; 9 = Sadness, 10 = Fear; 11 = Disgust. The following numbers represent pairwise mixed feelings coefficients: 12 = Amusement – Tenderness, 13 = Amusement - Anger, 14 = Amusement – Sadness, 15 = Amusement – Fear, 16 = Amusement – Disgust, 17 = Tenderness – Anger, 18 = Tenderness-Sadness, 19 = Tenderness –Fear, 20 = Tenderness-Disgust, 21 = Anger – Sadness, 22 = Anger – Fear, 23 = Anger-Disgust, 24 = Sadness-Fear, 25 = Sadness-Disgust, 26 = Fear – Disgust.

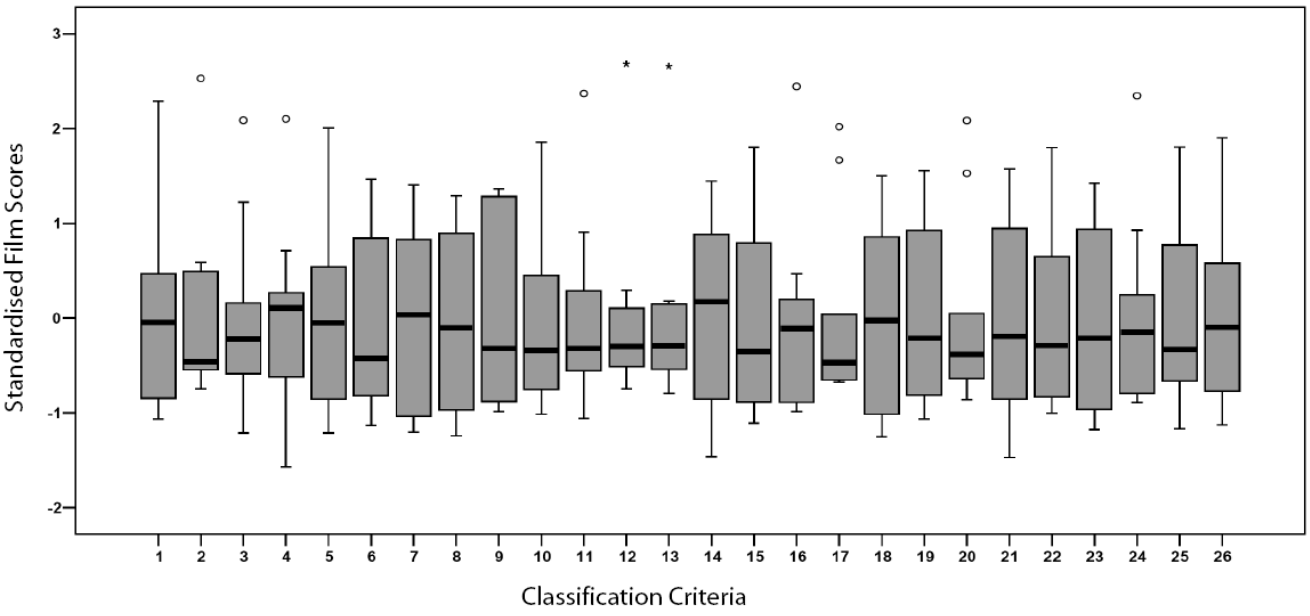

Arousal
8
7
6
5
4
3
2
1
0
ANG
FEAR
DISG
SAD
AM
TEND
NEUT


PA
3
2.5
2
1.5
1
0.5
ANG
FEAR
DISG
SAD
AM
TEND
NEUT


NA
3
2.5
2
1.5
1
0.5
ANG
FEAR
DISG
SAD
AM
TEND
NEUT


Positive composite score
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
ANG
FEAR
DISG
SAD
AM
TEND
NEUT


Negative composite score
4.5
4
3.5
3
2.5
2
1.5
1
0.5
ANG
FEAR
DISG
SAD
AM
TEND
NEUT


Standardised Film Scores
3
2
1
0
-1
-2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Classification Criteria

APPENDIX

Original French version of the modified Differential Emotions Scale used in the study:

English	French Translation
interested, concentrated, alert	attentif(ve), concentré(e), alerte
fearful, scared, afraid	apeuré(e), effrayé(e), terrifié(e)
anxious, tense, nervous	anxieux(se), tendu(e), nerveux(se)
moved	attendri(e)
angry, irritated, mad	en colère, irrité(e)
Ashamed, embarrassed	gêné(e), embarrassé(e), honteux(se)
warmhearted, gleeful, elated	heureux(se), exalté(e), épanoui(e)
joyful, amused, happy	joyeux(se), amusé(e), gai(e)
sad, downhearted, blue	triste, cafardeux(se)
satisfied, pleased	content(e), satisfait(e)
surprised, amazed, astonished	surpris(se), étonné(e), stupéfié(e)
loving, affectionate, friendly	amoureux(se), attiré(e), attaché(e)
guilty, remorseful	coupable, plein(e) de remords
disgusted, turned off, repulsed	dégoûté(e), écœuré(e), répugné(e)
disdainful, scornful, contemptuous	dédaigneux(se), méprisant(e)
calm, serene, relaxed	calme, serein(e), détendu(e)