

Emotion Elicitation Using Films

James J. Gross* and Robert W. Levenson

Psychology Department, University of California, Berkeley, USA

Researchers interested in emotion have long struggled with the problem of how to elicit emotional responses in the laboratory. In this article, we summarise five years of work to develop a set of films that reliably elicit each of eight emotional states (amusement, anger, contentment, disgust, fear, neutral, sadness, and surprise). After evaluating over 250 films, we showed selected film clips to an ethnically diverse sample of 494 English-speaking subjects. We then chose the two best films for each of the eight target emotions based on the intensity and discreteness of subjects' responses to each film. We found that our set of 16 films successfully elicited amusement, anger, contentment, disgust, sadness, surprise, a relatively neutral state, and, to a lesser extent, fear. We compare this set of films with another set recently described by Philippot (1993), and indicate that detailed instructions for creating our set of film stimuli will be provided on request.

INTRODUCTION

Because of its crucial role in normal and abnormal human functioning, social scientists have been increasingly eager to examine emotion under laboratory conditions. To do so, a number of emotion elicitation procedures have been used including: (a) interactions with trained confederates (e.g. Ax, 1953); (b) hypnosis (e.g. Bower, 1983); (c) repeating phrases (e.g. Velten, 1968); (d) facial muscle movements (e.g. Ekman, Levenson & Friesen, 1983); (e) imagery (e.g. Lang, 1979); (f) music (e.g. Sutherland, Newman, & Rachman, 1982); (g) slides (e.g. Lang, Ohman, & Vaitl, 1988; Wagner, 1990); and (h) films (e.g. Lazarus Speisman, Mordkoff, & Davison, 1962; McHugo, Smith, & Lanzetta, 1982; Philippot, 1993).

Among these methods, films have the desirable properties of being readily standardised, involving no deception, and being dynamic rather

Requests for reprints should be sent to Robert W. Levenson, Department of Psychology, University of California, Berkeley, CA 94720, USA.

* James J. Gross is now at the Department of Psychology, Stanford University.

This research was supported by NIMH National Research Service Award MH1003401 to the first author and NIA grant AG07476 and NIMH grant MH39895 to the second author.

than static. Films also have a relatively high degree of ecological validity, in so far as emotions are often evoked by dynamic visual and auditory stimuli that are external to the individual. One important limitation of the use of films, however, is that there has been no widely accepted set of emotion-eliciting film stimuli.

To address this issue, we began working five years ago to develop such a set of stimuli. In this article, we present data evaluating our most effective set of films (consisting of two films for eliciting each of eight emotional states: amusement, anger, contentment, disgust, fear, neutral, sadness, and surprise). We also compare our films with an independently developed set of films recently reported by Philippot (1993).

Previous Use of Films

The use of films in emotion research has a long history. Early "stress" studies often used films to elicit emotional reactions (e.g. Goodenough, Witkin, Koulack, & Cohen, 1975; Lazarus et al., 1962; Notarius & Levenson, 1979; Pillard, McNair, & Fisher, 1974). Here the issue was how to produce high intensity states of diffuse emotional arousal. In recent years, there has been increasing interest in studying more differentiated emotional states. Researchers working within a dimensional viewpoint (which holds that emotions represent points located on multiple dimensions, such as *valence/pleasantness* and *arousal/intensity*) have used films to elicit emotional states of a desired valence and intensity (e.g. Hubert & de Jong-Meyer, 1990; Van Rooijen & Vlaender, 1984). Researchers advocating a discrete emotions perspective (which holds that there are a finite number of distinct emotions that represent biologically based reactions that organise the individual's responses to important environmental events) have used films to elicit specific emotional states such as sadness and fear (e.g. Brown, Corriveau, & Monti, 1977; Engel, Frader, Barry, & Morrow, 1984; Marston, Hart, Hileman, & Faunce, 1984; Mewborn & Rogers, 1979).

Recently, in a study reflecting the discrete emotions perspective, Philippot (1993) assessed the efficacy of a set of 12 film clips (drawn from a pool of 20 candidate films) in eliciting six emotional states: amusement, anger, disgust, fear, neutral, sadness.¹ Sixty French-speaking Belgian students viewed these films and then reported on their emotional reactions using either: (a) a modified version of the Differential Emotions

¹ Philippot (1993) called his amusement films "happiness" films, although it is clear from the emotion self-reports he presented that these films elicited greater levels of amusement than happiness.

Scale (DES; Izard, Dougherty, Bloxom, & Kotsch, 1974); (b) a semantic differential scale; or (c) a free labelling format. Philippot reported success in eliciting differential emotion self-reports for amusement, sadness, and a neutral state, but had less success in eliciting anger, disgust, and fear.

Our work differs from Philippot's (1993) in a number of ways: (a) we aimed to find as many films as possible that elicited discrete emotional states in which one emotion predominated (Philippot's goal was to find pairs of films that produced equivalent differentiated emotional states, but not necessarily discrete emotions); (b) we considered a large pool of 250 films (rather than 20); (c) we employed a large ethnically diverse sample of 494 English-speaking subjects (rather than a small ethnically homogeneous sample of 60 French-speaking Belgian subjects); (d) we used a group-session format (rather than an individual-session format); (e) we attempted to elicit a larger set of emotions, including two positive emotions, amusement and contentment² (rather than just amusement), as well as surprise; and (f) we used a single self-report method based on 9-point Likert scales for each of 16 emotion terms (rather than three different self-report procedures). Fortunately, the similarity between our self-report procedure and Philippot's (1993) DES condition enables some comparison of the efficacy of the two sets of film stimuli. However, the aforementioned differences between the two studies suggest some caution in comparing findings.

METHOD

Subjects

A total of 494 undergraduates (229 men, 265 women) participated in group film-viewing sessions in order to fulfil a requirement of an introductory psychology course. The subjects were 17–43 years old (mean age = 19.3, SD = 1.7), and their ethnic identification approximated the demographics of the student population at the University of California, Berkeley (6% African-American, 42% Asian-American, 31% Caucasian, 10% Hispanic, and 11% Other).

² In our previous studies (e.g. Levenson, Ekman, & Friesen, 1990), happiness has often been the only positive affect studied. We now believe that it is important to differentiate between two major kinds of happiness—amusement and contentment.

Stimulus Films

Starting in 1988, we began soliciting our own research group, colleagues, film critics, video store employees, and film buffs for nominations of films that they thought would be effective elicitors of discrete emotions. These efforts produced a corpus of over 250 commercial films to which we added several film clips obtained from other investigators [Richard Davidson (University of Wisconsin, Madison), Paul Ekman (University of California, San Francisco), and Barbara Fredrickson (Duke University)]. Some of the nominations were full-length films, and from these we created short film clips by editing key sections.³ In addition, we generated one film comprised entirely of video test signals, which we thought might be affectively neutral.

From this large collection of film stimuli, 78 were selected for additional evaluation on the basis of: (a) length—films had to be relatively short; (b) intelligibility—the thematic content had to be understandable without additional explanation; and (c) discreteness—in our judgement, the film was *likely* to elicit a specific emotional state of either amusement, anger, contentment, disgust, fear, neutral, sadness, or surprise. Films in this set averaged 151 seconds in length (range = 8–1192 seconds) and most had sound tracks.

Procedure

The 78 films were shown to 31 groups of undergraduates (group size ranged from 3 to 30 subjects; mean = 16) on a 19-inch television monitor in a normal classroom. Each film was viewed by a minimum of 25 subjects (a minimum of 35 subjects viewed the 16 films that constitute our final set). Prior to viewing the films, subjects signed a consent form and answered several demographic questions. The experimenter stated that the purpose of the study was to learn more about emotion. Subjects were told that the films would be shown on a television monitor and that they should watch the films carefully, but could look away or shut their eyes if they found the films too distressing.

Subjects were shown approximately 10 films over the course of a one-hour session. Prior to each film, the experimenter stated that the screen would be blank for a while, and that subjects should use this time to “clear your mind of all thoughts, feelings, and memories”. The room lights were

³ The authors would like to thank the members of the film research team, including: Lisa Arnold, Christopher Gaines, Ronite Gluck, Teresa Goshgarian, Daniel Hadsall, Elizabeth Horn, Ila Kriplani, Jennifer Manly, Michelle Parra, Tina Ruiz, Claire Sauvageot, and Jon Wondrack

turned off, subjects were shown a blank screen for about 20 seconds, and then a film began. After each film, the room lights were turned on, and subjects were asked to complete a 16-item emotion self-report inventory (amusement, anger, arousal, confusion, contempt, contentment, disgust, embarrassment, fear, happiness, interest, pain, relief, sadness, surprise, and tension), which directed them to circle the number on the scale:

that best describes the *greatest amount* of each emotion you felt at any time during the film clip you have just seen. On this scale, 0 means you did *not feel even the slightest bit* of the emotion and 8 is the *most you have ever felt in your life*.

As noted above, this rating procedure, adapted from Ekman, Friesen, and Ancoli (1980), was quite similar to Philippot's (1993) DES condition. After completing the scale, subjects were asked whether they had seen the film previously.

Films were presented to groups of subjects with the following constraints: (a) to accustom subjects to the procedure, a film known to elicit low levels of contentment was shown first; (b) no two films targeting the same emotion were shown in a row; and (c) no more than three films of a particular valence (negative or positive) were shown consecutively.

RESULTS

Our analyses proceeded in several steps. We first conducted a series of analyses aimed at selecting our best films and at evaluating their efficacy in eliciting the eight target emotional states (amusement, anger, contentment, disgust, fear, neutral, sadness, surprise). We then compared our best films with Philippot's (1993) films using his published analyses and comparable analyses conducted on our films. Finally, we subjected our best films to additional analyses considering the role of sex, ethnicity, and prior viewing on subjects' reported emotions.⁴

⁴ In each of these analyses, we used the viewing of a film as the unit of analysis (see McHugo, Smith, & Lanzetta, 1982). To assess the possibility that there were effects attributable to the experimental session within which a given film was shown, we selected a subset of 16 films (see the section on selecting our best films) and conducted a Session × Emotion term (Amusement, Anger, Arousal, Confusion, Contempt, Contentment, Disgust, Embarrassment, Fear, Happiness, Interest, Pain, Relief, Sadness, Surprise, Tension) analysis of variance for each film. These analyses revealed Session or Session × Emotion effects for only 3 of the 16 films, suggesting a relatively minor influence of experimental session.

Selecting Our Best Films

Was the Target Emotion Elicited Most Strongly? For every film tested, we had an a priori intuition as to which emotion it would elicit most strongly. As a preliminary test, we examined the data to determine whether each film's anticipated target emotion had on average received a higher rating than the other six nontarget emotions. In most cases, our intuition was correct; for the few exceptions (4 of the 78 films), we treated the highest-rated emotion as the target emotion in subsequent analyses. To be considered a "neutral" film (i.e. a film that elicited very low levels of the 7 target emotions of interest to us), a film had to have mean emotion ratings of less than 2 points on the 9-point scale for each of the 7 target emotions. The final distribution of the 78 films among the 8 target emotional states was: amusement ($N = 16$), anger ($N = 8$), contentment ($N = 6$), disgust ($N = 16$), fear ($N = 13$), neutral ($N = 2$), sadness ($N = 11$), and surprise ($N = 6$).

Selecting the 2 Most Successful Films for Each Emotional State. Examining the mean emotion ratings, we were struck by the variability among these 78 film stimuli. Some of the films we had originally thought most likely to produce discrete emotional responses failed to do so; others

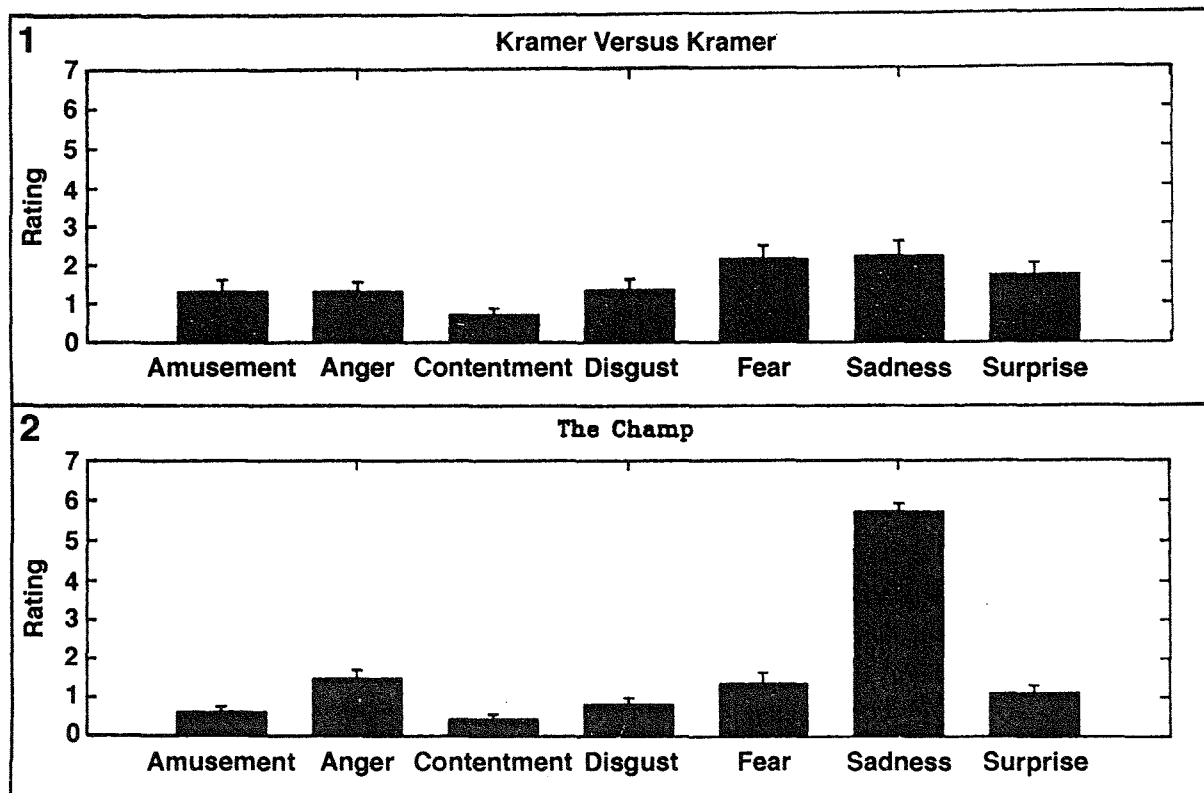


FIG. 1. Mean intensity of emotional ratings for two sadness films.

performed much better than we expected. To give some sense of how disparately different films performed, we present in Fig. 1 the results from two of our sadness films. The first was drawn from *Kramer Versus Kramer* (Benton, 1979), a scene in which a child falls and is rushed to a hospital room by his father, which many of our film nominators had recalled as being a real “tear-jerker”. The second was drawn from *The Champ* (Lovell & Zeffirelli, 1979), an arguably less familiar film in which a boy’s father dies after suffering a severe beating in the ring. As can be seen in Fig. 1, *Kramer Versus Kramer* produced levels of sadness that were not much greater than those for fear and surprise. *The Champ*, in contrast, produced levels of sadness that were much greater than those for any other emotion. In terms of ability to elicit a discrete state of sadness, visual examination of Fig. 1 suggests that *The Champ* was clearly more effective.

Rather than relying solely on this type of visual examination, we adopted a more objective criterion for selecting our most effective films which made use of both the *intensity* and the *discreteness* of the elicited response. We operationalised intensity as the mean level at which the target emotion was rated. We operationalised discreteness by deriving an idiosyncratic hit rate index (the percentage of subjects who indicated that they had felt the target emotion at least one point more intensely than any of the other six nontarget emotions).⁵ These measures of intensity and discreteness were combined into a “success” index by summing two *z*-scores, one derived by normalising the intensity scores for all of the films in that category, and the other derived by normalising the discreteness scores for all of the films in that category. We used this success index to select the two best films in each of the seven emotion categories, adding to them the two film stimuli that met the criterion for neutral emotion described earlier. These 16 films, which represent our recommendations for eliciting discrete emotions in the laboratory, are described in Table 1.

Evaluating the Efficacy of Our Best Films

Having identified our 16 best films, we next attempted to evaluate how effective they actually were in producing discrete emotions. We applied three tests:

⁵ These measures of intensity and discreteness were somewhat independent (the correlation between the two was 0.51). We also considered an alternative measure of discreteness—the rated intensity of the target emotion minus the mean rated intensity of the six nontarget emotions—but this measure proved to be quite redundant with the intensity measure ($r = 0.82$).

TABLE 1
Best 2 Films for Each Target Emotion

<i>Target Affect Film Name</i>	<i>Length of Film Clip (min:sec)</i>	<i>Mean Rating for Target Emotion (0-8 scale)</i>	<i>Hit Rate*</i>	<i>No. of Subjects Viewing Clip</i>
Amusement				
<i>When Harry Met Sally</i> (HARRY) Discussion of orgasm in cafe	2:35	5.54	93.1	72
<i>Robin Williams Live</i> (ROBIN) Comedy routine	7:59	5.86	84.1	63
Anger				
<i>My Bodyguard</i> (BODY) Bully scene	4:06	5.22	41.7	72
<i>Cry Freedom</i> (CRY) Police abuse protesters	2:36	6.05	22.0	59
Contentment				
<i>Waves</i> (WAVES) <i>Beach Scene</i> (BEACH)	0:58 0:23	3.46 3.58	58.1 43.2	42 343
Disgust				
<i>Pink Flamingos</i> (PINK) Person eats dog faeces	0:30	6.45	84.9	53
<i>Amputation</i> (AMP) Amputation of arm	1:03	5.60	80.3	125
Fear				
<i>The Shining</i> (SHINE) Boy playing in hallway	1:22	4.08	71.2	59
<i>Silence of the Lambs</i> (LAMBS) Basement chase scene	3:29	4.24	59.7	72
Neutral				
<i>Abstract Shapes</i> (SHAPE) <i>Colour Bars</i> (COLOUR)	3:25 1:30	– –	– –	55 35
Sadness				
<i>The Champ</i> (CHAMP) Boy cries at father's death	2:51	5.71	94.2	52
<i>Bambi</i> (BAMBI) Mother deer dies	2:19	5.35	76.4	72
Surprise				
<i>Capricorn One</i> (CAP) Agents burst through door	0:49	5.05	74.6	63
<i>Sea of Love</i> (SEA) Person startled by pigeons	0:09	4.22	66.7	54

*The hit rate is the percentage of subjects who indicated that they had felt the target emotion at least one point more intensely than any of the other six nontarget emotions.

1. *Discriminability*—using discriminant analysis, how well could the emotional state that was targeted by the 16 films be predicted from the self-report ratings?
2. *Discreteness*—using *t*-tests, was the rating for the target emotion term significantly greater than the other 15 emotions assessed?
3. *Similarity*—using cluster analysis and analysis of variance, how similar were the ratings for the two films for each target emotional state?

Discriminability. To assess the discriminability of the 16 best films, we conducted a discriminant function analysis in which subjects' ratings of the 16 emotion self-report terms were used to predict the emotion targeted by the films. This analysis showed fairly high levels of discrimination for each of the emotions: Amusement (82%), Anger (87%), Contentment (87%), Disgust (85%), Fear (78%), Neutral (66%), Sadness (86%), and Surprise (78%). Fewer than 7% of the film ratings in each of the non-neutral categories were categorised as neutral. Across all eight categories, the mean level of correct classification was 83%.

Discreteness. To assess discreteness, for each of the 16 films we conducted a 16-level within-subjects ANOVA (Emotion term: Amusement, Anger, Arousal, Confusion, Contempt, Contentment, Disgust, Embarrassment, Fear, Happiness, Interest, Pain, Relief, Sadness, Surprise, Tension). These analyses revealed significant main effects for Emotion term for each film. We then used *t*-tests to make pairwise comparisons between the target emotion and each non-target emotion. The results of these analyses are presented in Table 2.

These analyses provide further support of the efficacy of our best films. For amusement, disgust, sadness, and surprise, the target emotions were rated significantly higher than all of the other 15 nontarget emotions for both of the films. For anger, one film fully met this criterion and the other met it for all nontarget emotions except for disgust, for which ratings were not significantly different than for anger. For contentment, one film fully met the criterion and the other met it for all nontarget emotions except for happiness, which was rated at similar levels to contentment. Although neither of the fear films fully met the criterion, in both cases fear ratings were significantly higher than 13 of the nontarget emotions (in each case, fear ratings were not higher than interest or tension ratings).

Similarity. To assess similarity, we conducted a hierarchical cluster analysis on the 16 films using the ratings for the 16 emotion terms. Examination of the results of this analysis, which are presented in Fig. 2, reveals that at the eighth step the two films for each target affect form separate clusters. This indicates that each exemplar of a target emotion was

TABLE 2
Mean Emotion Intensity Ratings for Our Best 2 Films for Each Target Emotion

<i>Target Emotion Film Clip</i>	<i>Self-reported Emotion</i>																
	<i>F</i>	AMUS	ANGE	AROU	CFUS	CTEM	CTEN	DISG	EMBA	FEAR	HAPP	INTE	PAIN	RELI	SADN	SURP	TENS
<i>Amusement</i>																	
HARRY	127.50	5.54	0.31	2.96	0.36	0.43	2.63	0.44	2.29	0.30	3.35	3.99	0.21	0.97	0.15	1.54	0.89
ROBIN	101.99	5.86	0.29	2.52	0.74	0.68	3.42	0.71	0.65	0.08	4.62	4.65	0.29	1.71	0.16	2.08	0.78
<i>Anger</i>																	
BODY	89.93	0.90	5.22	2.08	1.68	4.60	0.75	4.96 ^a	0.82	2.07	0.55	3.50	2.51	0.30	3.74	1.54	4.10
CRY	74.32	0.39	6.05	2.71	2.59	4.97	0.69	5.49	1.14	3.42	0.46	3.56	4.58	0.32	5.42	2.59	4.61
<i>Contentment</i>																	
WAVES	23.52	1.50	0.19	1.29	0.48	0.46	3.46	0.12	0.05	0.43	2.88	2.14	0.07	1.64	0.67	0.60	0.69
BEACH	312.59	2.74	0.30	1.23	0.89	0.48	3.58	0.23	0.17	0.29	3.53 ^a	3.27	0.29	2.14	0.66	1.32	0.63
<i>Disgust</i>																	
PINK	63.99	2.43	0.70	1.17	1.91	1.83	0.51	6.45	1.06	0.43	0.43	1.62	0.51	0.43	0.56	3.49	0.98
AMP	108.03	0.74	0.59	2.02	2.19	0.86	0.33	5.60	0.34	1.75	0.16	2.62	2.86	0.33	0.70	1.94	3.50
<i>Fear</i>																	
SHINE	83.43	1.05	0.36	2.00	2.31	0.33	0.72	0.15	0.08	4.08	0.49	4.17 ^a	0.34	0.27	0.37	1.34	4.37 ^a
LAMBS	56.88	1.74	1.22	2.53	1.28	0.97	1.25	2.07	0.38	4.24	1.07	4.49 ^a	0.79	0.50	0.63	2.03	4.74
<i>Sadness</i>																	
CHAMP	67.11	0.62	1.50	1.46	1.46	0.65	0.44	0.83	0.44	1.37	0.27	3.13	3.69	0.56	5.71	1.13	2.88
BAMBI	48.67	1.69	2.65	1.39	0.62	1.94	1.12	1.69	0.44	1.78	0.97	2.86	2.83	0.64	5.35	1.27	2.21

<i>Surprise</i>																	
CAP	72.37	0.81	0.38	1.02	3.84	0.48	0.73	0.42	0.10	2.65	0.30	2.90	0.35	0.41	0.43	5.05	3.03
SEA	53.28	1.44	0.22	1.52	1.61	0.24	0.57	0.24	0.33	2.94	0.65	2.74	0.28	0.72	0.17	4.22	2.83
<i>Neutral</i>																	
SHAPE	—	0.91	1.07	0.61	2.49	0.84	0.98	0.55	0.16	0.27	0.76	0.98	0.38	0.45	0.25	0.80	1.22
COLOUR	—	1.00	1.03	0.70	2.65	0.94	1.09	0.53	0.15	0.06	0.62	0.88	0.09	0.37	0.09	1.26	1.00

Note. Numbers in each row that have superscripts do not differ reliably from the target emotion for that row (which is in italics) at the $P < 0.05$ significance level.

Column key: AMUS = Amusement, ANGE = Anger, AROU = Arousal, CFUS = Confusion, CTEM = Contempt, CTEN = Contentment, DISG = Disgust, EMBA = Embarrassment, FEAR = Fear, HAPP = Happiness, INTE = Interest, PAIN = Pain, RELI = Relief, SADN = Sadness, SURP = Surprise, TENS = Tension.

Row key: HARRY = *When Harry Met Sally* (Reiner, Scheinman, Stolt, & Nicolaidis, 1989), ROBIN = *Robin Williams Live* (Morra, Brezner, & Gowers, 1986), BODY = *My Bodyguard* (Devlin & Bill, 1980), CRY = *Cry Freedom* (Spencer, Briley, & Attenborough, 1987), WAVES = *Waves* (Ekman), BEACH = *Beach Scene* (Ekman), PINK = *Pink Flamingos* (Waters, 1973), AMP = *Amputation* (Ekman), SHINE = *The Shining* (Kubrick, 1980), LAMBS = *Silence of the Lambs* (Saxon, Utt, Bozman, & Demme, 1991), CHAMP = *The Champ* (Lovell & Zeffirelli, 1979), BAMBI = *Bambi* (Disney & Hand, 1942), CAP = *Capricorn One* (Lazarus & Hyams, 1978), SEA = *Sea of Love* (Bregman, Stroller, & Becker, 1989), SHAPE = *Abstract Shapes* (ScreenPeace screensaver), COLOUR = *Colour Bars* (Gross).

more similar to the other exemplar of that emotion than to any of the films for the other emotions.

Comparing Our Films with Philippot's Films

Despite the differences between methods and goals described earlier, we thought it might be useful for researchers interested in eliciting discrete emotions for us to compare the efficacy of our films with Philippot's (1993).

Our approach to this comparison was dictated in part by the analyses that he reported and by several methodological differences between the two studies. Because Philippot did not develop films to elicit contentment or surprise, we limited our comparisons to films that elicited the remaining six emotional states (amusement, anger, disgust, fear, neutral, and sadness). In these comparisons, we used the 10 DES categories employed by Philippot in his DES condition ($N = 20$) and the comparable rating categories from our study.⁶ Because his scale anchoring ("very strongly") was different than our own ("most in your life"), we focused on the relative discreteness of the emotion self-reports rather than comparing the absolute intensity of the target emotions. Because he presented his emotion self-report data only after collapsing across the two films that targeted each emotion, we did the same.

Discriminability. We first compared how well subjects' emotion self-reports distinguished films that targeted one emotion from films that targeted another emotion. To do so, we followed Philippot's procedure of conducting discriminant function analyses for six target emotional states (amusement, anger, disgust, fear, neutral, and sadness). For Philippot's film set, we used the results he reported (Philippot, 1993) from the discriminant analyses using the 10 emotion terms in his DES condition. For our films, we conducted an identical analysis using the same 10 emotion terms from our data. Results from his and our film sets are presented in Table 3.⁷

⁶ The equivalence between our emotion terms and the DES was as follows (in each set our term appears to the left of the equal sign and the DES terms appear to the right): Amusement = amused, joyful, merry; Anger = angry, irritated, mad; Contempt = disdainful, scornful, contemptuous; Disgust = disgusted, turned off, repulsed; Fear = fearful, scared, afraid; Happiness = warmhearted, gleeful, elated; Interest = interested, concentrated, alert; Sadness = sad, downhearted, blue; Surprise = surprised, amazed, astonished; Tension = anxious, tense, nervous.

⁷ To reassure ourselves that these results were not an artefact of the discrepancy between our and Philippot's sample sizes (each film was viewed by at least 35 subjects in our study compared to 10 subjects in Philippot's study), we randomly selected 10 subjects per film (for

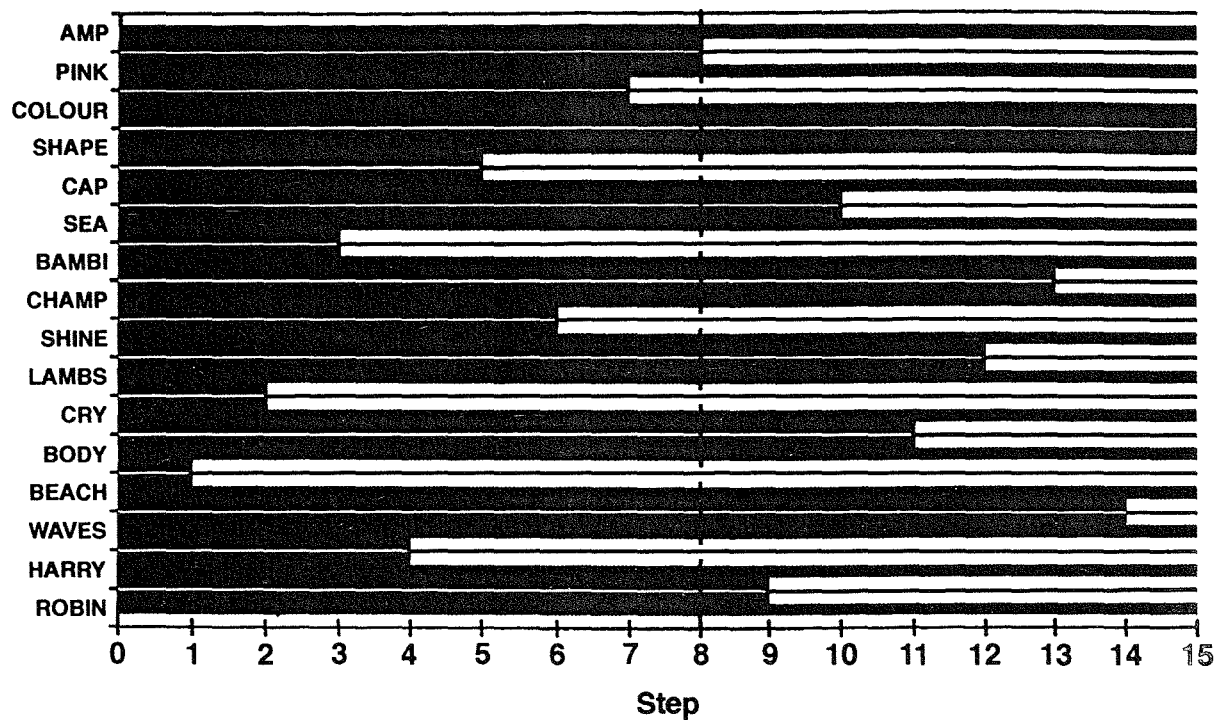


FIG. 2. Hierarchical cluster analysis of our 16 best films. AMP = *Amputation*; PINK = *Pink Flamingos*; COLOUR = *Colour Bars*; SHAPE = *Abstract Shapes*; CAP = *Capricorn One*; SEA = *Sea of Love*; BAMBI = *Bambi*; CHAMP = *The Champ*; SHINE = *The Shining*; LAMBS = *Silence of the Lambs*; CRY = *Cry Freedom*; BODY = *My Bodyguard*; BEACH = *Beach Scene*; WAVES = *Waves*; HARRY = *When Harry Met Sally*; ROBIN = *Robin Williams Live*.

Examination of Table 3 reveals that the overall discrimination among the films in our sample (86%) was significantly greater than it was in Philippot's sample (70%), $Z = 4.00$, $P < 0.001$. Table 3 also indicates that, among specific emotions, our amusement, anger, and neutral films were significantly more discriminable than his corresponding films.

Discreteness. We next compared the extent to which pairs of films produced reports of the target emotion that were greater than those of other emotions. To do so we followed Philippot's procedure of conducting within-subjects *t*-tests to determine whether mean levels of the target emotion were greater than each of the nine non-target emotions. In this analysis, we used the same five target emotion categories as he did. In addition, because his sample size was smaller than ours, to achieve comparable statistical power we matched sample sizes by randomly selecting 10 of our subjects for each of our 10 best films (two films for each of five target emotion categories used by Philippot). For Philippot's film set,

times, each time drawing a new random sample. The average number of times films were correctly categorised over these 25 random samples was very similar to the result reported in the text.

TABLE 3
Percentage of Correct Classifications across 6 Target Emotions

<i>Target</i>	<i>Our Films</i>	<i>Philippot Films</i>	<i>Z</i>	<i>P</i>
Amusement	0.91	0.76	1.80	0.034
Anger	0.84	0.62	2.04	0.019
Disgust	0.84	0.75	1.03	0.152
Fear	0.83	0.75	0.86	0.198
Neutral	0.91	0.62	3.09	0.001
Sadness	0.85	0.71	1.49	0.066
Total	0.86	0.70	4.00	0.001

TABLE 4
Mean Emotion Intensity Ratings for Our Films (Random Sample of 10 Subjects per Film) and for Philippot's Films in his DES Condition (10 Subjects per Film)

<i>Self-reported Emotion</i>	<i>Target Emotion in Film</i>									
	Amusement		Anger		Disgust		Fear		Sadness	
	A*	B**	A	B	A	B	A	B	A	B
Amusement	5.60	2.86	0.75	0.05	1.28	0.20	0.95	0.37	1.00	0.51
Anger	0.35	0.14	5.35	2.81	1.22	3.05 ^b	0.70	0.44	1.95	0.70
Contempt	0.30	0.48	4.30	1.23	2.06	2.15	0.50	0.29	1.11	0.15
Disgust	0.40	0.38	5.50 ^a	2.90 ^b	6.28	3.75	0.55	1.17	1.10	0.29
Fear	0.05	0.14	2.10	1.97 ^b	2.06	2.15	3.35	2.09	1.55	0.14
Happiness	3.85	1.47	0.40	0.00	0.22	0.10	1.25	0.22	0.65	0.32
Interest	4.05	3.00 ^{***}	2.90	3.65 ^{***}	2.17	3.60 ^{***}	4.55 ^a	3.57 ^{***}	3.15	3.54 ^{***}
Sadness	0.40	0.09	4.60 ^a	2.14 ^b	1.06	1.60	0.35	0.55	5.80	2.16
Surprise	1.55	0.44	1.80	1.27	2.72	2.45 ^b	1.40	1.30	1.58	0.14
Tension	0.90	0.51	3.55	2.54 ^b	3.17	2.65 ^b	3.90 ^a	2.67 ^b	2.65	1.01

Note. Compare cell means vertically. Numbers in each column that have superscripts do not differ reliably from the target emotion for that column (which is in italics) at the $P < 0.05$ significance level.

* Column A is our sample of randomly selected film viewing episodes (10 for each film). Subjects reported their emotional responses using a 9-point Likert scale, 0 = none, 8 = the most in my life.

** Column B is Philippot's (1993) sample of film viewing episodes (10 subjects watched each film). Subjects reported their emotional responses using a 5-point Likert scale (0 = not at all, 4 = very strongly).

*** Philippot did not indicate whether self-reports of the target emotion were greater than self-reports of interest.

we used the results he reported (Philippot, 1993) from his DES condition. For our films, we conducted identical analyses using the same 10 emotion terms from our data. Results from his and our film sets are presented in Table 4.

Examination of Table 4 reveals that for four of the five emotion categories (amusement, anger, disgust, and sadness) our films elicited more discrete emotional states than did Philippot's. Whereas his *amusement* films elicited levels of amusement that were not higher than levels of interest, our amusement films elicited more amusement than any other emotion. His *anger* films elicited levels of anger that were not greater than levels of disgust, fear, interest, sadness, and tension. Our anger films elicited higher levels of anger than any other emotion except disgust and sadness. His *disgust* films elicited levels of disgust that were not higher than levels of anger, interest, surprise, and tension. Our disgust films elicited higher levels of disgust than any other emotion. His *sadness* films elicited levels of sadness that were not higher than levels of interest. Our sadness films elicited higher levels of sadness than any other emotion. For the fifth emotion category, fear, both sets performed similarly; both his and our *fear* films elicited levels of fear that were not greater than levels of interest and tension.

Similarity. As noted earlier, we differed from Philippot in terms of stated goals for the film stimuli. Whereas our primary interest was in finding films that elicited discrete emotional states, Philippot wanted to find pairs of films that elicited equivalent emotional states, even if these were not discrete emotional states. To test the comparability of his two exemplar films for each emotion, Philippot (1993) reported a $2 \times 6 \times 10$ (Exemplar \times Target Emotion \times Emotion Term) ANOVA. To provide a point of comparison, we computed a similar ANOVA on our data, matching his smaller sample size with a randomly selected set of subjects so that both analyses would have similar statistical power. Whereas Philippot found no significant main or interaction effects involving Exemplar, in our data the Exemplar \times Target Emotion \times Emotion Term interaction was significant, $F(45,500) = 2.10, P < 0.001$. Although failure to reject the null hypothesis surely does not demonstrate the degree of equivalency between pairs of films, the results from these parallel analyses suggest that Philippot's pairs of films might be more closely matched than were ours.

Emotional Analyses of Our Films: Effects of Sex, Ethnicity, and Prior Viewing

To provide an estimate of the overall effects of sex, ethnicity, and prior viewing, we computed the average intensity of the target emotion for our 14 films that targeted nonneutral emotion. A two-level (Women, Men)

ANOVA revealed that women reported greater levels of target emotions than did men [mean target emotion: Women = 4.98, Men = 4.44, $F(1,1195) = 19.06$, $P < 0.001$]. A four-level (African-American, Asian-American, Caucasian, Hispanic) ANOVA failed to reveal an Ethnicity effect, $F(3,1080) = 0.40$, n.s. A two-level (Seen previously, Not seen previously) ANOVA revealed that subjects who had seen the film previously reported greater levels of target emotions than did subjects who had not seen the film previously [mean target emotion: Seen previously = 5.10, Not seen previously = 4.64, $F(1,1188) = 9.67$, $P = 0.002$].

DISCUSSION

In 1981, Polivy reviewed the available methods for eliciting emotion in the laboratory and expressed extreme pessimism about the possibility of eliciting discrete emotions under laboratory conditions. More than a dozen years later, we find ourselves considerably more optimistic. Our findings (and to some extent those of Philippot, 1993), indicate that films can be found that meet fairly stringent criteria for the elicitation of amusement, anger, contentment, disgust, a relatively neutral state, sadness, surprise, and, to a lesser extent, fear. This suggests that carefully selected films may be useful in solving the perennial problem of how to elicit discrete emotions in the laboratory, and that comparable neutral stimuli may be used to control for the effects of simply watching a film.⁸

Differences Among Emotions

This is not to say, however, that all emotions are equally easy to elicit using films. Even with our best films, there is considerable variability in the discreteness and intensity of the emotional responses across target emotion categories. In terms of discreteness, idiographic hit rate percentages range from the eighties for amusement, disgust, and sadness (meaning that these films met our criteria for successfully eliciting the target emotion in more than 80% of these subjects), to an average of 32% for the anger films. This variability in hit rates underlines the importance of carefully verifying the success of films (or any other emotion elicitation procedure) in producing the targeted emotional reactions in individual subjects.

⁸ We sought "neutral" films that would elicit little or none of seven target emotions. For our purposes, we were not unduly concerned by the slight elevations in confusion self-reports produced by these films (which were likely engendered by showing such nonemotional films during experimental sessions explicitly described as concerned with emotion). In different applications, other kinds of "neutral" films may be preferable.

It is also important to consider differences in intensity of emotional response. Our films fell into five groups, with disgust being elicited the most intensely; followed by amusement, anger, and sadness; followed by surprise; followed by fear; followed by contentment (see Table 2). It is important to recognise that even carefully selected films of similar lengths can elicit quite different intensities of emotional responses. Failure to assess and to consider such differences might lead to erroneous conclusions (e.g. attributing found effects to differences between emotions, when they are in fact due to differences in intensity).

The More Difficult Emotions: Anger, Contentment, and Fear

Some emotions seem to be more difficult to elicit than others using films; for example, we only found one successful film for anger and contentment, and did not find any successful films for fear. Why has it been so difficult to elicit discrete responses for these emotions?

Selection of film stimuli. Perhaps our candidate films for these three emotions were poorly chosen. However, given the extensive efforts we took to find adequate stimulus materials and similar difficulties reported by Philippot (1993) in eliciting anger and fear (he did not attempt to elicit contentment), we do not believe this is a complete explanation.

Selection of emotion terms. Perhaps the 16 emotion terms in our self-report instrument unevenly sampled the universe of emotion terms. For example, if anger, fear, and contentment were presented along with near synonyms, it would be difficult for them to meet our criterion for discreteness.

For contentment, we think this might be a valid concern. The contentment film (*Beach Scene*) that did not meet criterion failed because it also produced elevated levels of the near synonym "happiness" (see Table 2). By our criterion, comparable elevation of self-reported happiness along with contentment denoted failure. However, if these emotions were considered to be synonymous or hierarchical (i.e. contentment is a subtype of the superordinate category of happiness), then what we considered failure to elicit discrete contentment might be considered a success. Viewed in this light, contentment would join amusement, disgust, sadness, neutral, and surprise, constituting a set of six emotions for which we have found *two* successful film elicitors.

For anger and fear, however, this possibility seems less viable, given that our 16-item lexicon did not include more terms that were semantically close than for any of the other negative emotions. In fact, disgust, for which we found two quite successful films, arguably had a much closer term to

compete with (i.e. contempt) than either anger or fear. How, then, are we to account for our difficulties in eliciting anger and fear?

Anger. For anger, one film (*My Bodyguard*) failed to meet our criterion because the levels of anger that it produced were not greater than the levels of disgust (see Table 2). The other anger film (*Cry Freedom*) did meet our criterion, but examination of the response profile (Table 2) indicates that a number of other emotions were more elevated than we would have liked. Although we may have simply chosen poor anger-eliciting films, we are becoming increasingly convinced that elicitation of discrete anger with brief films is going to be extremely difficult, if not impossible. We know from other work that discrete anger can be elicited using a more social, interpersonal stimulus (e.g. attempts by spouses to resolve marital problems: Gottman & Levenson, 1992). With films, it appears that there is a natural tendency for anger to co-occur with other negative emotions (McHugo et al., 1982; Philippot, 1993; Tourangeau & Ellsworth, 1979).

Fear. For fear, neither of our two best films met our criterion; for both, levels of fear were not greater than those of interest or tension. In fact, for *The Silence of the Lambs*, the level of tension was actually greater than the level of fear (see Table 2). Again, it may be that we have not yet located those films that will effectively elicit discrete fear. However, we expect that this confluence of fear, tension, and interest may be a natural one, a view that is supported by Philippot's (1993) finding identical elevations of these three emotion terms in response to his fear films. Fear engenders vigilance to the immediate environment, which is likely to be reflected in elevated ratings of interest. Fear also begets a state of preparation for action and of motoric readiness, which may be reflected in elevated ratings of tension.

Caveats and Concerns

There are several caveats and concerns that must be raised concerning the findings presented in this article, including: (a) what we mean by "discreteness;" (b) the relationship between emotion self-reports and other aspects of the emotional response; (c) demand characteristics; and (d) the role of individual differences.

Defining "discreteness". Whether we are judged to have succeeded in eliciting discrete emotions using films clearly depends on what one means by "discrete". If one requires high levels of a single emotion with *no traces* of any other emotion, we suspect discrete emotional responses would be very hard to find (either in day-to-day living or inside the laboratory). On the other hand, if one merely requires that a subject feel more of a certain emotion than one or two others, then discrete emotions likely are quite commonplace. In our work, we have taken an intermediate position, saying

that a discrete emotional response has occurred when a subject reports feeling a single emotion more intensely than 15 other positive and negative emotions. Given that a large part of our research agenda involves trying to understand how emotions differ from each other, this definition has seemed most useful to us. For other purposes, other definitions might prevail.

The use of self-report measures. All of the data presented are based on subjects' self-reports. As we have stated previously (Levenson et al., 1990), we do not believe that self-report is the *sine qua non* of emotion. Nonetheless, we do believe that finding a set of films that produce the desired profile of self-reported emotion is a good starting point, providing a foundation for determining whether such films also produce behavioural and physiological signs of the target emotion.

Demand characteristics. Demand characteristics were clearly present in our studies. We told subjects that we were studying emotion and that we would like them to rate their emotional responses to the films they reviewed. Thus, it was transparent that we were hoping that the films would elicit emotion, and it is possible that subjects responded with the emotion they thought the film *was intended* to produce, even though they felt no emotion themselves. Although we do not believe this to be the case—our instructions clearly indicated that subjects were to describe how *they felt* during the film—additional data will be required to address this legitimate concern.

Individual differences. It appears that individual differences are an important aspect of the emotional responses to films and to other stimuli. In this work, we have provided a preliminary test of how several sources of individual variation moderate the emotional impact of films. Our finding of more intense reports by women than men is consistent with other findings that women often report more intense emotional experiences than men (LaFrance & Banaji, 1991; Shields, 1991). Our finding of no ethnic differences in film report suggests that these films may be useful with a wide variety of populations. Our finding that prior viewing is associated with more intense reports suggests a sensitisation effect, likely due in part to a greater understanding of the total emotional contact of the film from which the film clip was taken. In the future, it will be important to consider the role of other kinds of individual differences (e.g. personality variables) in order to understand better the sizeable individual variability in emotional responses to film stimuli.

Recommendations and Future Directions

When we started our research five years ago, there were no empirically validated sets of film stimuli designed to elicit a broad range of discrete emotions. With our work and that of Philippot (1993), there are now at

least two such sets. Philippot's set may be most useful for those who require pairs of films that produce equivalent emotional states. Our set of 16 films (presented in Table 1) represent our best suggestions for emotion researchers who are interested in using films to elicit discrete emotions.

Our set of films should not be seen as definitive. We regard the development of film stimuli as an ongoing process. We fully expect that we and others will eventually find films that are more effective in eliciting discrete emotions than those presented here. In the meantime, we hope that the presentation of extensive analyses of the efficacy of our films will facilitate comparison between these and other films. Also, recognising that we have thus far only sought films that would elicit the eight emotional states of amusement, anger, contentment, disgust, fear, neutral, sadness, and surprise, we hope that others will make similar efforts for other important emotions such as contempt, pride, guilt, and embarrassment.

Obtaining a Set of Our 16 Best Films

We are committed to sharing the fruits of our labours with others working in this area. Given that most of the film clips in our set were extracted from commercial films, we cannot provide copies of the films themselves. All of the commercial films in the set, however, are currently available on videotape, and, upon request, we will provide the editing instructions needed to produce the same film clips that we used. For the non-commercial film clips, we will indicate how to make or obtain these stimuli as well.

Manuscript received 20 October 1993
Revised manuscript received 5 April 1994

REFERENCES

- Ax, A.F. (1953). The physiological differentiation between fear and anger in humans. *Psychosomatic Medicine*, *15*, 433–442.
- Benton, R. (Producer and Director) (1979). *Kramer Versus Kramer* [Film]. CA: Columbia/TriStar Home Video, Burbank, CA.
- Bower, G. (1983). Affect and cognition. *Philosophical Transactions of the Royal Society of London*, *302B*, 387–402.
- Bregman, M., Stroller, L. (Producers), & Becker, H. (Director) (1989). *Sea of Love* [Film]. MCA/Universal Home Video, Los Angeles, CA.
- Brown, W.A., Corriveau, D.P., & Monti, P.M. (1977). Anger arousal by a motion picture: A methodological note. *American Journal of Psychiatry*, *134*, 930–931.
- Devlin, D. (Producer), & Bill, T. (Director). (1980). *My Bodyguard* [Film]. Fox Hills Video, Culver City, CA.
- Disney, W. (Producer), & Hand, D. (Director) (1942). *Bambi* [Film]. Walt Disney Home Video, Burbank, CA.

- Ekman, P., Friesen, W.V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of Personality and Social Psychology*, 39, 1124–1134.
- Ekman, P., Levenson, R.W., & Friesen, W.V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221, 1208–1210.
- Engel, G.L., Frader, M., Barry, C., & Morrow, G.R. (1984). Sadness evoked by a film on grief: An experimental study. *International Journal of Psychiatry in Medicine*, 14, 1–30.
- Goodenough, D.R., Witkin, H.A., Koulack, D., & Cohen, H. (1975). The effects of stress films on dream affect and on respiration and eye-movement activity during rapid-eye-movement sleep. *Psychophysiology*, 12, 313–320.
- Gottman, J.M., & Levenson, R.W. (1992). Marital processes predictive of later dissolution: Behavior, physiology and health. *Journal of Personality and Social Psychology*, 63, 221–233.
- Hubert, W., & de Jong-Meyer, R. (1990). Psychophysiological response patterns to positive and negative film stimuli. *Biological Psychology*, 31, 73–93.
- 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.
- Kubrick, S. (Producer and Director) (1980). *The Shining* [Film]. Warner Home Video, Burbank, CA.
- LaFrance, M., & Banaji, M. (1992). Toward a reconsideration of the gender-emotion relationship. In M. Clark (Ed.), *Review of personality and social psychology* (pp. 178–201). Beverly Hills, CA: Sage.
- Lang, P.J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16, 495–512.
- Lang, P.J., Ohman, A., & Vaitl, D. (1988). *The international affective picture system* (photographic slides). Center for Research in Psychophysiology, University of Florida: Gainesville, FL.
- Lazarus, P. (Producer), & Hyams, P. (Director) (1978). *Capricorn One* [Film]. CBS/Fox Video, New York.
- Lazarus, R.S., Speisman, J.C., Mordkoff, A.M., & Davison, L.A. (1962). A laboratory study of psychological stress produced by a motion picture film. *Psychological Monographs*, 76 (whole number 553).
- Levenson, R.W., Ekman, P., & Friesen, W.V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, 27, 363–384.
- Lovell, D. (Producer), & Zeffirelli, F. (Director) (1979). *The Champ* [Film]. MGM/Pathé Home Video, Culver City, CA.
- Marston, A., Hart, J., Hileman, C., & Faunce, W. (1984). Toward the laboratory study of sadness and crying. *American Journal of Psychology*, 97, 127–131.
- McHugo, G.J., Smith, C.A., & Lanzetta, J.T. (1982). The structure of self-reports of emotional responses to film segments. *Motivation and Emotion*, 6, 365–385.
- Mewborn, C.R., & Rogers, R.W. (1979). Effects of threatening and reassuring components of fear appeals on physiological and verbal measures of emotion and attitudes. *Journal of Experimental Social Psychology*, 15, 242–253.
- Morra, B., Brezner, L. (Producers), & Gowers, B. (Director) (1986). *Robin Williams Live* [Film]. Vestron Video, Stamford, CT.
- Notarius, C.I., & Levenson, R.W. (1979). Expressive tendencies and physiological response to stress. *Journal of Personality and Social Psychology*, 37, 1204–1210.
- Philippot, P. (1993). Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition and Emotion*, 7, 171–193.
- Pillard, R.C., McNair, D.M., & Fisher, S. (1974). Does marijuana enhance experimentally induced anxiety? *Psychopharmacologia*, 40, 205–210.

- Polivy, J. (1981). On the induction of emotion in the laboratory: Discrete moods or multiple affect states? *Journal of Personality and Social Psychology*, 41, 803–817.
- Reiner, R. (Producer and Director), Scheinman, A., Stolt, J., & Nicolaides, S. (Producers) (1989). *When Harry Met Sally* [Film]. New Line Home Video, Los Angeles, CA.
- Saxon, E., Utt, K., Bozman, R. (Producers), & Demme, J. (Director) (1991). *Silence of the Lambs* [Film]. Orion Pictures, New York, NY.
- Shields, S.A. (1991). Gender in the psychology of emotion: A selective research review. In K.T. Strongman (Ed.), *International review of studies on emotion* (Vol. 1, pp. 227–245). New York: Wiley.
- Spencer, N., Briley, J. (Producers), & Attenborough, R. (Director) (1987). *Cry Freedom* [Film]. MCA/Universal Home Video, Los Angeles, CA.
- Sutherland, G., Newman, B., & Rachman, S. (1982). Experimental investigations of the relationship between mood and intrusive unwanted cognitions. *British Journal of Medical Psychology*, 55, 127–138.
- Tourangeau, R., & Ellsworth, P.C. (1979). The role of facial response in the experience of emotion. *Journal of Personality and Social Psychology*, 37, 1519–1531.
- Van Rooijen, L., & Vlaander, G.P. (1984). Dramatic induction of depressive mood. *Journal of Clinical Psychology*, 40, 1318–1322.
- Velten, E. (1968). A laboratory task for the induction of mood states. *Behaviour Research and Therapy*, 6, 473–482.
- Wagner, H.L. (1990). The spontaneous facial expression of differential positive and negative emotions. *Motivation and Emotion*, 14, 27–43.
- Waters, J. (Director) (1973). *Pink Flamingos* [Film]. Lightning Video, Stamford, CT.