

Reisenzein, R., & Hofmann, T. (1993). [Discriminating emotions from appraisal-relevant situational information: Baseline data for structural models of cognitive appraisals.](#) *Cognition and Emotion*, 7, 271-293.

<http://www.tandf.co.uk/journals/pp/02699931.html>

Discriminating Emotions from Appraisal-relevant Situational Information: Baseline Data for Structural Models of Cognitive Appraisals

Rainer Reisenzein

Free University, Berlin, Germany

Thomas Hofmann

University of Technology, Berlin, Germany

Cognitive emotion theorists assume that the quality of emotions is determined by the appraisal of the eliciting states of affairs. Accordingly, a central criterion for the evaluation of structural models of cognitive appraisal is their capacity to discriminate between emotions on the basis of the proposed appraisal dimensions. It is suggested that a good model should approximate subjects' "natural" ability to distinguish emotions on the basis of appraisal-relevant situational information. Corresponding data for 23 common emotions, which can serve as a baseline for the evaluation of cognitive appraisal theories, are reported, and various factors that may have deflated the discrimination rates obtained so far in empirical studies are discussed.

INTRODUCTION

A major concern of recent cognitively oriented research on emotions consists of the enumeration of the basic dimensions of cognitive appraisal and their associations with different emotions. Several structural models of cognitive appraisal have been proposed during recent years (e.g. Dalkvist & Rollenhagen, 1989; Frijda, 1986; Ortony, Clore, & Collins, 1988; Roseman, 1984; Scherer, 1984; Smith & Ellsworth, 1985; Solomon, 1976; Weiner, 1986). These models differ with regard to the number and identity of the dimensions of appraisal that are proposed, the assumptions (if such are spelled out) concerning the association between particular emotions and patterns of appraisal, and with regard to the range of emotions that each theory seeks to cover (see e.g. Roseman, Spindel, & Jose, 1990). Attempts to test these models presume criteria of empirical adequacy by which they can be evaluated. These criteria concern both the aptness of a model when considered separately, and when compared with other theories.

Recent empirical research has particularly emphasised one criterion of adequacy, namely the degree to which an appraisal model permits the discrimination among different emotions (e.g. Ellsworth & Smith, 1988; Frijda, 1987; Frijda, Kuipers, & ter Schure, 1989; Gehm & Scherer, 1988; Manstead & Tetlock, 1989; Smith & Ellsworth, 1985, 1987; Smolenaars & Schutzelaars, 1986/87; Tesser, 1990). Judged by this criterion, most of the

theories tested so far can be said to have received some degree of empirical support, inasmuch as appreciable associations between emotions and the postulated dimensions of cognitive appraisal were obtained. To illustrate, Smith and Ellsworth (1985) found that knowledge of subjects' ratings of emotional episodes for 15 common affects on six appraisal dimensions permitted the correct statistical classification of 42% of these situations (37% if corrected for chance; Cohen's [1960] $\kappa \times 100$). Similarly, in the most extensive investigation to date, Frijda et al. (1989), using as predictors ratings on 19 (study 1) and 23 (study 2) appraisal scales designed to tap several further appraisal dimensions in addition to those proposed by Smith and Ellsworth, obtained statistical classification rates of 36% (chance-corrected 34%; study 1) and 43% (chance-corrected 41%; study 2) for the 32 emotions that were included.

It is more difficult to say: (1) how strong this support can be considered to be in an *absolute* sense (i.e. does a discrimination rate of 30-40% mean that the model is good, moderate, or poor?); and (2) what the discrimination data tell about the *relative* merits of the different theories. Considering first the last-mentioned issue, a direct comparison of most of the existing studies is difficult because they differed, apart from the appraisal dimensions included, with regard to several other factors. These include the number and kind of emotions that were investigated, the "reality mode" of the emotion-eliciting situations that were judged (e.g. real, remembered or hypothetical situations), the measurement of appraisals and emotions, and the statistical methods used to determine the predictive capacity of the proposed dimensions (e.g. discriminant analysis vs. multiple regression). A way out of this problem are comparative model-testing studies in which these potentially confounding factors are controlled and the discriminatory capacity of different appraisal dimensions, and combinations of such dimensions, is systematically examined (e.g. Roseman et al., 1990). As to the first-mentioned issue, whether one judges the empirically obtained association between appraisals and emotions as high or low depends on how strong one expects this relation to be on theoretical, methodological, and empirical grounds. Concerning theory, most authors seem to assume that each distinct emotion is "associated", in some way or other, with a distinct pattern of appraisal. If this assumption is correct, perfect discrimination between emotions on the basis of appraisals should theoretically be possible, and the comparatively low empirical discrimination scores obtained so far could be interpreted as indicating that important dimensions of appraisals are missing in current models. But some cognitive emotion theorists (e.g. Frijda, 1986; Smolenaars & Schutzelaars, 1986/87) have suggested that part of the everyday distinctions between emotions reflect differences in aspects of emotions other than appraisals (such as action tendencies or bodily feedback).¹ If so, the association between appraisals and emotions obtained in empirical studies can a priori be expected to be less than perfect, perhaps substantially so (cf. Smolenaars & Schutzelaars, 1986/87). Furthermore, discrimination between emotions

¹ This assumption presupposes that bodily feelings, action tendencies, etc. are not themselves completely determined by cognitive appraisals. Frijda et al. (1989, p. 225) suggest, for example, that emotion-related action tendencies may in part also depend on the individual's emotional response propensities, such as his or her proneness to fear, anger, or joy.

should be improved if information about the further relevant aspects of emotions is also taken into account. In line with this prediction, Frijda et al. (1989) found that the additional inclusion of action tendencies ("action readiness modes") as predictors increased the percentage of correctly classified emotions to up to 59% (Frijda et al., 1989, study 2).

However, even if a one-to-one association between emotions and patterns of appraisal is assumed to exist, empirical associations substantially stronger than those obtained so far need not necessarily be expected. First, some of the (presumably) different emotions included in the studies may, in fact, not be very sharply distinguished by people. It has frequently been suggested that the demarcations drawn by the average person between many of the mental states denoted by different emotion terms are fuzzy or even nonexistent and that emotion language is therefore full of "quasisynonyms" (e.g. Marx, 1982; Osgood, 1966; Russell, 1980; Tomkins & McCarter, 1964). Related to this point, there may be substantial interindividual differences in the use of some emotion words. Secondly, methodological factors could have deflated the truly existing associations between emotions and appraisal. Emotion-eliciting events frequently evoke multiple emotions (cf. Smith & Ellsworth, 1987); because the appraisal ratings typically referred to the total situation, the resulting "mixtures" of appraisals may have blurred the association between emotions and appraisals (Frijda et al., 1989; see also Reisenzein & Hofmann, 1990; Roseman et al., 1990). In addition, the scales used to assess the appraisal dimensions may have partly suffered from low reliability or validity (e.g. Roseman et al., 1990; Tesser, 1990); and the statistical prediction models used may not have been fully adequate for the task (Frijda et al., 1989; Tesser, 1990).

Hence, the implications of the obtained discrimination scores for the aptness of the tested appraisal models remain ambiguous. They could indicate that important appraisal dimensions have been omitted, that some of the distinctions between emotions depend crucially on nonappraisal elements, that some of the examined emotions were not clearly distinguished by people or that there were substantial inter-individual differences in the use of some emotion words, that methodological problems were responsible, or that a combination of these factors was at work. Ideally, one would like to compare the statistical discrimination scores with empirically obtained baseline data on people's ability to discriminate between emotions on the basis of information about cognitive appraisals, analogous to the strategy used in research on the components of facial expression that people use to infer emotional states (e.g. Ekman, 1982). There, the aptness of a proposed "component model" can be evaluated by comparing the predictive success of that model with data on the recognition of emotions from facial expressions by naive judges. However, in contrast to facial expressions of emotions which can simply be photographed and then presented to judges, cognitive appraisals, being mental states, are not directly accessible to observation. In fact, it could be argued that one has to first know the dimensions of appraisal to be able to determine people's ability to discriminate emotions on their basis, implying that this ability cannot be determined in a theory-independent way at all. However, although it may not be possible to determine people's ability to distinguish between emotions on the basis of appraisals in a way that does not already presuppose a structural theory of appraisals, there is a

related issue that *can* be examined in a theory-independent manner and that provides information which is nearly as useful. This is the question of how well people can distinguish between emotions on the basis of *appraisal-relevant situational information*.² Provided that this information is sufficiently detailed to permit reliable inference of the appraisals of the emotion experiencer, data on people's ability to distinguish emotions on the basis of such information can be used as a baseline for the evaluation of appraisal models. The reason is that cognitive appraisal theorists assume that, whatever effects situations have on emotions, these effects are mediated (at least for more complex emotions, such as those considered here) by cognitive appraisals (e.g. Lazarus, 1982). If this mediational assumption is correct, "situations should add little or . . . nothing to prediction when the effect of dimensions is controlled" (Tesser, 1990, p. 218). Hence—unless this basic assumption is abandoned—it follows that whatever discrimination between emotions is possible on the basis of appraisal-relevant situational information must in some way be "picked up" or explained by appraisal models.

The aim of the studies reported later is to provide such discrimination data for 23 common emotions. Our approach was based on the assumption that salient perceived features of emotion-eliciting events—including those features that would be regarded as (contents of) appraisals by the theorists mentioned earlier (e.g. the valence of an event, its causal locus, or its probability)—are usually accessible to introspection and can therefore be reported. Indeed, when people in everyday life communicate to one another about their emotions, they seem frequently to refer either to their appraisals or to appraisal-relevant situational information to help the other person understand why they felt in a particular way in that situation (see Fiebler, 1990; Laucken, 1989). Our method may be regarded as a laboratory version of this procedure. Subjects were interviewed for reports of emotion-eliciting events for 23 common emotions. Special instructions were used to maximise the chances that their subjective appraisals of the eliciting situations, or at least appraisal-antecedent situational information from which the appraisals could be reliably inferred, rather than mere "objective" reports of what had happened, were elicited. Subsequently, different subjects were presented with condensed descriptions of these accounts, in which all information other than about eliciting conditions and appraisals was eliminated, and were asked to infer the emotion likely to be experienced by the story protagonists. In study 1, they had to do this by selecting the emotion terms deemed appropriate from the list of 23 terms, whereas in studies 2 and 3, they had to classify a subset of the emotion scenarios either into 23 standard categories which served as markers (study 2), or to sort them freely according to perceived similarity of the feelings elicited by these situations (study 3).

² The term "appraisal-relevant situational information" was suggested to us by one of the reviewers. As explained later, this information consists, more precisely, of a mixture of (selective) information about appraisals and about situational features from which appraisals can potentially be inferred.

STUDY 1

Method

Subjects

Altogether, 78 subjects participated in study 1, 27 of them in the initial interview designed to obtain exemplars of emotion-eliciting situation descriptions, and the remaining 51 in the subsequent emotion discrimination experiment (main study). All subjects were students at the Free University Berlin who responded to posted advertisements in which paid volunteers for an unspecified "psychological experiment" were sought. The subjects were paid DM 10 (interview) and DM 40 (discrimination study) respectively for participation. Three subjects did not return their questionnaires, reducing the final sample size of the main study to 47 (including 1 subject who completed only two-thirds of the questionnaire). Of the 27 subjects of the interview study, 12 were male and 15 female (age $M = 24.2$, $sd = 4.3$); of the 47 subjects from the main experiment whose data were used, 25 were male and 22 female (age $M = 23.8$, $sd = 3.2$).

Selection of Emotion Types

A total of 23 emotions were included for study. Their selection was guided by both theoretical and empirical criteria. That is, on the one hand we wanted to include affects which have been accorded central importance by contemporary emotion theorists of both cognitive (e.g. Roseman, 1984; Weiner, 1986) and other (e.g. Izard, 1977; Plutchik, 1980) provenance; on the other hand, states which are regarded by lay subjects as typical examples of the category "emotion" as determined through empirical studies (e.g. Averill, 1975; Davitz, 1969; Fehr & Russell, 1984; Marx, 1982; Schmidt-Atzert, 1981) were to be included. The resulting list (cf. Table 1), which is a compromise between these two criteria, seems to be a fairly representative sample of the emotional spectrum (cf. Shaver, Schwartz, Kirson, & O'Connor, 1987).

Procedure

Interview. The 27 subjects were interviewed individually by one of the two authors; about half by each. The subjects were asked to recount, for each of the 23 emotions, an episode (preferably one that had recently happened) where they had experienced the respective affect as the dominant emotion. In order to elicit, as much as possible, the subjects' cognitive appraisals of the emotional events, or at least appraisal-relevant antecedent information from which the appraisals could be inferred, the subjects were instructed to describe the events from their personal perspective, that is, as they had viewed them at the time when they occurred, and to provide enough detail to enable the experimenter—a person unfamiliar with their general life circumstances, attitudes, preferences, goals, etc.—to understand why they had felt the particular emotion in the respective situation. If necessary, this instruction was reiterated during the interview until the experimenter had the impression that he understood why the subject had felt the emotion in the situation reported (however, no report was ever

rejected). In addition, the experimenter attempted to determine at which moment during an unfolding situation the affect first occurred and, if the situation was a complex one involving several emotions, which aspect of the situation, in the subject's opinion, was responsible for the affect. The interview was tape-recorded with the subjects' permission (no one declined) and later transcribed.

Questionnaire Construction. A total of 544 descriptions of emotional situations were obtained from the 27 subjects ($M = 20.2$, $sd = 2.5$). From this material we intended to construct, for each emotion, 20 non-redundant descriptions of eliciting events for use in the subsequent discrimination experiment. In a first step, the transcribed reports were condensed into short first-person stories according to the following rules. (1) First, background information concerning the subject's study area, sex, general life circumstances, preferences and attitudes toward specific issues, etc. was provided, if such information had been indicated by the subject to be relevant for an understanding of the emotional reactions. Examples are: "I am a passionate chess player"; "my relationship with my boyfriend has been going badly during the past few months". (2) This information was followed by the description of the event immediately responsible for the subject's emotion. The unfolding event (as appraised by the subject) was described up to and including the moment at which the emotion first occurred, ending with the standard phrase "At this moment, I experienced the following emotion: _____", or a semantically equivalent sentence. (3) All emotion terms as well as metaphors referring to emotions, if such had been mentioned by the subjects, were eliminated from the descriptions; similarly, if a subject had reported physiological symptoms, expressive or instrumental reactions, or responses of social partners, these kinds of information were also discarded. This elimination procedure precludes the possibility that subjects in the main study could have inferred emotions from information other than about appraisals and situational cues. In a second step, we eliminated from each emotion category those situations that seemed redundant. If an affect category still contained more than 20 stories after this elimination procedure, as many as necessary were randomly removed to reduce their number to 20. For six of the emotion categories, fewer than 20 situations remained after elimination of redundant stories. Therefore, several additional subjects were interviewed specifically for these affects until 20 non-redundant situation descriptions had been obtained.

Finally, the 23 (affects) \times 20 (number of situations per affect) = 460 stories were randomised and typed as a questionnaire. The length of the stories varied from a minimum of 12 words to a maximum of 120 words, with most of the stories being from 40 to 70 words long. To illustrate what the stories looked like, here is an example of a better recognised story from the embarrassment category:

I am not accustomed to being the focus of other people's attention.—When I finished school, a big celebration was organised. I was the second best of my class. Although I did not regard this as a particularly noteworthy achievement (standards in our class were low), everybody made a big thing of it. During the celebration, I was asked to step out in front where everybody told me how great I was. When I was standing there, being the focus of

everybody's attention, and heard everybody tell me how great it was, I experienced the following emotion:

(This story was judged as an example of embarrassment by 78% of the subjects in the subsequent discrimination experiment.)

Discrimination Study. Based on the results of pretesting, the questionnaire was divided into three parts which were to be completed during three separate 2 hour sessions, distributed over a period of 4 weeks. At each session, one-third of the questionnaire was handed out to the 51 subjects together with a list of the 23 emotion names arranged in alphabetical order. The subjects' task was to infer which emotions the person from whom the stories had been obtained might have experienced in each scenario, by selecting the emotion name which seemed to be the most appropriate label (*first answer*). In addition, the subjects were allowed to select a further emotion name (*second answer*) as a second-best guess. For each subject, the order in which the stories appeared in the questionnaire was additionally randomised through random arrangement of the 121 pages of the questionnaire.

Results³

A total of 21,323 first and 5927 second answers were obtained from the 47 subjects included in the data analysis. Missing first responses were uniformly due to subjects who overlooked a page or two of the questionnaire.

The relatively low percentage of second responses (on average, for 28% of the stories) may have been due to the subjects' relatively great certainty regarding their first choices, but other factors, such as the desire to shorten the laborious task, were probably also influential.

To answer the question of how well the subjects discriminated among the 23 emotions on the basis of appraisal-relevant situational information, we computed the chance-corrected proportion of agreement (Cohen's [1960] kappa) between their responses and the emotion labellings of the scenarios by the encoders (i.e. the persons from whom the scenarios had been obtained). In addition to an overall agreement coefficient, kappa coefficients were also computed separately for the 23 emotion categories (conditional kappa; see Fleiss, 1971; Hubert, 1977). Furthermore, all agreement statistics were calculated once considering only the subjects' first answers, and again when taking both their first and second answers into account. In the latter case, an agreement was defined to be present if at least one of the two answers corresponded to the comparison standard, and the chance correction factor involved in the computation of kappa was appropriately adjusted to reflect the increased probability of chance agreement. The conditional kappa values for the 23 emotion categories, their ranks, and the corresponding raw proportions are shown in Table 1 in columns [1] to [6]. As is to be expected, given the large number of categories and the equal distribution of stories across the categories, the kappa values are only slightly lower than the corresponding raw proportions.

³ A more detailed description of the results, including additional data, is available from the first author on request.

Considering first the agreements of the subjects' first responses with the encoders' (reported) emotional state, it can be seen from column [2] of Table 1 that the conditional kappa values ranged from a minimum of 0.30 (remorse/regret) to a maximum of 0.88 (disgust/revulsion) with a mean of 0.64 (sd = 0.15), which coincided with the overall agreement coefficient calculated for the pooled data. Three emotion categories had kappas \geq 0.80, 9 \geq 0.70, 14 \geq 0.60, and 20 \geq 0.50. Even for the three emotions which had agreement scores $<$ 0.50 (contempt, shame, and remorse/regret), the conditional kappas were far beyond the chance level. Furthermore, the correct response category was always the most frequently chosen one and, for all emotions except remorse/regret, it was significantly ($P < 0.05$; binomial z-test) more likely to be selected than any other category. On average, five times as many answers fell within the correct response category than within the most frequently chosen incorrect one (i.e. the latter contained, on average, only 13% of the responses). Hence, it is evident that the subjects discriminated rather well between the emotions. Examination of columns [4] and [5] of Table 1 reveals that the agreement coefficients uniformly increased when both first and second answers were taken into account, whereas the rank order of the coefficients remained nearly constant (Spearman rho = 0.99). The maximum kappa value is now raised to 0.91 (disgust/revulsion) and the minimal value (remorse/regret) to 0.36, with a mean of 0.70 (sd = 0.13). One affect now had a kappa value \geq 0.90; eight had kappas \geq 0.80; 12 \geq 0.70; 18 \geq 0.60, and 21 \geq 0.50. Hence, even if the subjects missed the correct emotion category at their first guess, they often had it ready as their second-best answer.

TABLE 1
Study 1: Overall Agreement of Subjects' Responses with the Criterion and Significant Disagreements

| Emotion Category | Agreement with Criterion | | | | | | Categories where Significant Disagreements Occurred ^a |
|--|--------------------------|-------|------|-------------------|-------|------|---|
| | 1st Answer | | | 1st or 2nd Answer | | | |
| | Rank | Kappa | Prop | Rank | Kappa | Prop | |
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] |
| Anger/Rage | 9 | 0.72 | 0.74 | 9 | 0.79 | 0.81 | disappointment (0.10); contempt (0.08) |
| Anxiety/Fear | 6 | 0.76 | 0.77 | 7.5 | 0.80 | 0.81 | |
| Contempt | 21 | 0.49 | 0.51 | 20 | 0.58 | 0.61 | disgust (0.14); anger (0.13); disappointment (0.09); pity (0.07) |
| Disappointment | 10 | 0.69 | 0.71 | 10 | 0.78 | 0.80 | anger (0.09) |
| Discontentment/ Dissatisfaction with self | 14 | 0.60 | 0.62 | 13 | 0.69 | 0.72 | hopelessness (0.11); anger (0.08) |
| Disgust/Revulsion | 1 | 0.88 | 0.88 | 1 | 0.91 | 0.92 | |
| Embarrassment | 17.5 | 0.53 | 0.55 | 18 | 0.60 | 0.63 | shame (0.11); discontentment (0.06) |
| Envy | 11 | 0.67 | 0.68 | 12 | 0.71 | 0.73 | jealousy (0.06); discontentment (0.06) |
| Gratitude | 6 | 0.76 | 0.77 | 5.5 | 0.82 | 0.83 | joy (0.10); relief (0.07) |
| | | | | | | | |
| Guilt | 17.5 | 0.53 | 0.55 | 17 | 0.62 | 0.64 | remorse (0.20); shame (0.09) |
| Hope | 2 | 0.84 | 0.85 | 2 | 0.88 | 0.89 | joy (0.08) |
| Hopelessness/Resignation | 15 | 0.57 | 0.59 | 16 | 0.63 | 0.65 | disappointment (0.17); sadness (0.07) |
| Jealousy | 13 | 0.61 | 0.62 | 14 | 0.67 | 0.68 | envy (0.11); disappointment (0.07) |
| Joy/Happiness | 12 | 0.66 | 0.69 | 11 | 0.76 | 0.79 | pride (0.09); surprise (0.06) |
| Loneliness | 8 | 0.75 | 0.76 | 7.5 | 0.80 | 0.81 | |
| Love | 19.5 | 0.50 | 0.52 | 19 | 0.59 | 0.61 | joy (0.26); hope (0.06) |
| Pity/Sympathy | 4 | 0.78 | 0.79 | 3.5 | 0.82 | 0.83 | sadness (0.06) |
| Pride | 6 | 0.76 | 0.77 | 5 | 0.82 | 0.83 | joy (0.16) |
| Relief | 3 | 0.80 | 0.81 | 3.5 | 0.83 | 0.84 | joy (0.06) |
| Remorse/Regret | 23 | 0.30 | 0.32 | 23 | 0.36 | 0.39 | discontentment (0.29); anger (0.12); disappointment (0.09); guilt (0.06) |
| Sadness/Sorrow | 16 | 0.56 | 0.58 | 15 | 0.65 | 0.67 | disappointment (0.17); loneliness (0.08) |
| Shame | 22 | 0.38 | 0.40 | 22 | 0.47 | 0.50 | embarrassment (0.23); discontentment (0.09); remorse (0.07); guilt (0.07) |
| Surprise | 19.5 | 0.50 | 0.51 | 21 | 0.56 | 0.58 | joy (0.24) |
| Mean | | 0.64 | 0.65 | | 0.70 | 0.72 | |
| sd | | 0.15 | 0.14 | | 0.13 | 0.13 | |

^aNumbers in parentheses are proportion disagreements.

An analysis of the observed *disagreements* is shown in column [7] of Table 1, which contains, for each of the emotions, those disagreements which occurred with conditional probabilities significantly different from the base probabilities of the respective response category in the total pool of observations (as determined by the binomial z-test with alpha set at 0.05; see e.g. Allison & Liker, 1982). As can be seen, systematic disagreements occurred for all emotions except anxiety/fear, disgust/revulsion, and loneliness, but these disagreements were in all cases restricted to a small subset of the possible 22 "incorrect" affects, mostly to those which intuitively appear to be related to the respective "correct" emotions (see also Shaver et al., 1987). The systematic disagreements accounted for 55% of all observed disagreements; therefore only 16% of the total first responses of the subjects remain as random answers. Within the 23 emotion categories, random disagreements ranged from 6% (contempt) to 28% (embarrassment). Furthermore, emotions with lower agreement scores tended to

be confused with greater numbers of affects ($r = -0.55$).

To obtain more information about the sources of the differing overall agreement scores of the various emotions and the observed disagreements, analyses analogous to the ones just reported were performed separately for the 20 scenarios within each emotion category. The main results of these analyses were as follows. (1) *All* emotion categories contained at least some well-discriminated scenarios. If only well-discriminated scenarios were considered, high accuracy scores were obtained for nearly all emotions. For example, if only the best-discriminated situation from each category was considered, agreement ranged from 1.0 (anxiety/fear, disgust/revulsion, gratitude, hope, and loneliness) to 0.70 (remorse/regret) with $M = 0.91$ and $sd = 0.09$; and 14 emotions had proportion agreement values ≥ 0.90 , 21 ≥ 0.80 , and all 23 ≥ 0.70 . (Data for the 10 best-discriminated scenarios are presented later in Table 2.) (2) Every single one of the 460 situations had at least one mode (meaning, in the present context, a significant elevation from baseline), implying that no situation was responded to completely randomly. (3) Of the 460 situations, 229 (49.8%) were unimodal, 179 (38.9%) were bimodal, 47 (10.2%) had three, and 5 (1.1%) had four modes. Hence, about half of the scenarios suggested more than a single emotion as the dominant one to the decoders. (4) Finally, for 73 (or 16%) of the scenarios, the most frequently used response category was different from the emotion intended by the encoder. These scenarios had uniformly low agreement scores (≤ 0.40). Thus, a substantial number of cases were observed which suggested a different emotion to the decoders than the one intended by the encoder. However, all but three of these scenarios were also multimodal, and in only 11 of these cases did the intended emotion not appear as an additional significant mode. If the *modal* response to these scenarios is regarded as the correct response category, the overall agreement kappa increased to 0.67 (first response only) and 0.74 (first or second response correct).

Discussion

Considering the potential methodological limitations of the procedure, which could have negatively affected the results (in particular, that not all of the appraisal-relevant situational information was elicited in the interview), the findings suggest that subjects can rather well discriminate between emotions, at least those considered, on the basis of appraisal-relevant situational information. The overall discrimination performance of our naive subjects was, for example, not much below the reliability generally considered sufficient in observational or judgement studies using trained observers (e.g. Rosenthal, 1982), and for many of the emotion categories, it was entirely adequate. In addition, it appears that nearly perfect discrimination ($M = 0.91$) can be obtained if complex or ambiguous scenarios are excluded.

Nevertheless, the finding that rather low average agreement scores were obtained for some emotions raises the possibility that recognition of these affects may crucially depend on information other than about cognitive appraisals (e.g. Frijda et al., 1989; Smolenaars & Schutzelaars, 1986/87). However, the fact that even these emotion categories contained at least a few well-discriminated scenarios speaks against this possibility and suggests that other factors may have been responsible. First, the lower accuracy

rates may have been due, in part, to linguistic factors; that is, there may be relatively great inter-individual inconsistencies concerning the meanings of the corresponding emotion labels. This possibility was further examined in study 2. Secondly, the less well-discriminated scenarios of these emotion categories may have been primarily situations which, rather than being ambiguous (i.e. situations which left it unclear which of two or more emotions occurred) were complex, that is, unambiguously suggested the simultaneous presence of several emotions (cf. Smith & Ellsworth, 1987). In the latter case, the subjects' uncertainty and possible errors (in the present study) would have been due only to the fact that they could not clearly decide which emotion was the *dominant* one. To check this latter possibility, we presented the 231 multimodal situations to two additional subjects together with the modal responses and asked them to indicate for each of these situations how likely it was, in their opinion, that the respective emotions had been experienced by the story protagonist (on a scale from 1 = very unlikely to 7 = practically certain). For nearly all of the scenarios, the raters were relatively certain (score > 5) that *all* modal affects had occurred. Thus, complexity of the eliciting situations may indeed have been a major factor **responsible for lower agreement on particular emotions**.

STUDIES 2 AND 3

To validate the results of study 1, two further studies were conducted. Both were partial replications of study 1, using different methods designed to overcome some of the potential limitations of the first investigation. In particular, we attempted to control for the possibility that the moderate agreement scores for some emotions observed in study 1 were due to the fact that the subjects disagreed to some extent on the meanings of the corresponding emotion words. To assess the importance of this factor, a fixed and free sorting method was used in studies 2 and 3, respectively. Linguistic differences concerning the emotion terms should be controlled by this method to some degree, because scenarios should be grouped into the same target category if they are seen as eliciting the same emotion(s) as the target situation, regardless of how the subjects choose to call them. Due to space reasons, and because the results of these studies were very similar to those of study 1, they will be described only briefly (see footnote 3).

Subjects and Procedure

Subjects were 21 (study 2) and 24 (study 3) students from the same subject pool as those of study 1. In study 2, the subjects were presented with 23 "standard scenarios" consisting of the best-discriminated scenario from each of the 23 emotion categories (according to the results of study 1) and were asked to sort the better-discriminated half of the remaining scenarios, which had been typed on index cards, into the 23 categories according to perceived similarity to the emotion(s) experienced in the standard situations. After the subjects had completed the sorting, they were asked to name the emotion(s) believed to be experienced in the 23 categories of situations, using from one to three emotion terms of their choice. Subsequently, they were presented with the list of the 23 emotion

names used in study 1 and were again asked to label the categories, this time using only the names appearing on the list.

Because of the high cognitive demands of the free-sorting procedure, in study 3 only 46 scenarios, 2 out of each of the 23 emotion categories (deck A), had to be sorted. However, this task was replicated for a second set of 46 scenarios (deck B). Following standard instructions for free-sorting tasks (e.g. Fillenbaum & Rapoport, 1971; Miller, 1969), the subjects were asked to sort the scenarios into piles on the basis of perceived similarity of emotions. They were allowed to use as many piles as they wanted and to put as many scenarios in a pile as they wished. Subsequently, they were asked to suggest one or more labels for the emotions experienced in each cluster of scenarios.

Results and Discussion

Study 2

The major results of study 2 are presented in Table 2. Agreement with the standard, averaged across emotion categories, was $\kappa = 0.73$ ($sd = 0.17$). This is slightly lower than the corresponding statistic obtained in study 1 for the same scenarios ($M = 0.77$, $sd = 0.14$). Five emotions (envy, guilt, jealousy, regret/remorse, and shame) were discriminated significantly better in study 2 than in study 1 (as determined by t -tests on the means of the individual kappa scores; $\alpha = 0.05$, $df = 66$), whereas the reverse was true for six different emotions (anger, anxiety, disappointment, embarrassment, joy, and surprise); the remaining 12 did not differ.

The increases in classification accuracy observed in study 2 (which included regret/remorse and guilt, the two affects which had the lowest accuracy scores in study 1) would be consistent with the hypothesis that the moderate to low scores obtained for these emotions in study 1 were in part due to inter-individual differences concerning the meanings of the emotion labels. In contrast, the observed decrements in accuracy were unexpected, but closer examination suggested that they were either due to the fact that the single scenarios, which served as markers, suggested as a salient emotion one that was not available as a response in study 1 (e.g. the surprise scenario also suggested admiration and fascination), or that idiosyncratic features of the criterion scenarios, which are not generally characteristic of the situations associated with the target emotion, were also attended to in classification. This explanation was confirmed by additional analyses (cf. footnote 3).

Columns [6] and [7] of Table 2 show that on average 79% ($sd = 21$) of the self-produced emotion labels per category were literally correct, (column [6]), ranging from 38% (surprise) to 100% (anxiety/fear). Of those labels that were not literally correct, about one-third may be considered to be sufficiently similar to the preassigned labels to be taken, in the present context, as semantic equivalents of them (these labels are italicised in column [7]). If these labels are included among the correct ones, on average 86% of the labels are correct ($sd = 16$). The results of the additional category labelling, where the emotion names had to be selected from the list of labels used in study 1, were even more impressive: For all but three emotions (embarrassment, hopelessness, and surprise) this labelling was 100% correct. In sum, the results of study 1 were largely

TABLE 2
Study 2: Agreement Indices, Significant Disagreements and Labelling of the Categories

| Emotion Category | Agreement with Criterion | | | | Categories where Significant Disagreements occurred ^b | Self-produced Emotion Labels | |
|--|--------------------------|-------|----------------------|-------|--|------------------------------|--|
| | Study 2 | | Study 1 ^a | | | Literally correct | Other |
| | Rank | Kappa | Rank | Kappa | | | |
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] |
| Anger/Rage | 15 | 0.70 | 10 | 0.84 | | 0.95 (20) | sore(1) |
| Anxiety/Fear | 10.5 | 0.80 | 4 | 0.89 | | 1.00 (21) | |
| Contempt | 16 | 0.65 | 21 | 0.58 | Anger (0.08); Disgust (0.07) | 0.48 (10) | superiority (7); lack of understanding for other (3); hate (1) |
| Disappointment | 19.5 | 0.56 | 11 | 0.81 | Hopeless (0.11); Anger (0.09); Sad (0.06) | 0.95 (20) | lack of solidarity (1) |
| Discontentment/ Dissatisfaction with self | 17.5 | 0.63 | 15.5 | 0.72 | Hopeless (0.10) | 0.48 (10) | own insufficiency (5); feeling useless (1); frustra- tion about self (2); disappointed about self (2); tired (1) |
| Disgust/Revulsion | 2 | 0.94 | 1 | 0.96 | | 0.95 (20) | lack of understanding for other (1) |
| Embarrassment | 21 | 0.54 | 19.5 | 0.65 | Shame (0.26) | 0.33 (7) | perplexed (3); insecure (9); incompetent (1); forsaken (1) |
| Envy | 3 | 0.92 | 12 | 0.80 | | 0.90 (19) | jealous of other's possessions (1); jealous (1) |
| Gratitude | 8 | 0.84 | 4 | 0.89 | | 0.86 (18) | loyalty (1); liking (1); needy (1) |
| Guilt | 10.5 | 0.80 | 17 | 0.68 | Remorse (0.15) | 0.90 (19) | self-reproachful (1); remorse (1) |
| Hope | 4.5 | 0.90 | 2 | 0.93 | | 0.86 (18) | optimism (1); new beginning (2) |
| Hopelessness/Resignation | 13 | 0.75 | 13 | 0.78 | Dissatisfied (0.14) | 0.57 (12) | despairing (2); discouraged (1); lack of control (1); depressed (1); stressed (1); frustrated (3) |
| Jealousy | 6.5 | 0.86 | 15.5 | 0.72 | | 0.95 (20) | feeling ignored (1) |
| Joy/Happiness | 22 | 0.53 | 9 | 0.85 | Love (0.26); Surprise (0.07) | 0.95 (20) | joyful surprise (1) |
| Loneliness | 4.5 | 0.90 | 7.5 | 0.87 | | 0.95 (20) | loss (1) |
| Love | 12 | 0.76 | 18 | 0.67 | Joy (0.07); Gratitude (0.06); Surprise (0.06) | 0.57 (12) | liking (1); elated (2); happiness (6) |
| Pity/Sympathy | 6.5 | 0.86 | 7.5 | 0.87 | | 0.90 (19) | empathy (1); disgust (1) |
| Pride | 1 | 0.95 | 4 | 0.89 | | 0.81 (17) | self-enhancement (2); joy at one's success (1); happy (1) |
| Relief | 9 | 0.82 | 6 | 0.88 | | 0.95 (20) | liking for other (1) |
| Remorse/Regret | 17.5 | 0.63 | 22.5 | 0.46 | Guilt (0.11); Dissatisfied (0.09) | 0.71 (15) | self-reproachful (1); guilty (2); unfairness (1); angry at self (2) |
| Sadness/Sorrow | 14 | 0.71 | 14 | 0.77 | Disappointment (0.22) | 0.76 (16) | feeling of loss (5) |
| Shame | 19.5 | 0.56 | 22.5 | 0.46 | | 0.86 (18) | embarrassment (3) |
| Surprise | 23 | 0.21 | 19.5 | 0.65 | Joy (0.46) | 0.38 (8) | struck (1); funny (1); fascinated (3); admiration (5); authority (1); autonomous (1); unclassifiable (1) |
| Mean | | 0.73 | | 0.77 | | 0.79 (16.5) | |
| sd | | 0.17 | | 0.14 | | 0.21 (4.4) | |

^aFor the 230 scenarios used in study 2.

^bNumbers in parentheses are proportion disagreements significant at alpha = 0.05.

replicated, and some indication was obtained that linguistic factors (disagreement on the meanings of emotion terms) contributed to the lower accuracy scores obtained in study 1 for some of the emotions.

Study 3

On average, 15.4 (sd = 4.1) clusters were constructed for deck A, and 16.2 (sd = 3.9) for deck B. For decks A and B separately, individual incidence matrices were first formed, which were then summed across subjects. These matrices were correlated with a theoretically derived "perfect" discrimination matrix, which would have been obtained if all subjects had formed 23 clusters containing exactly the two scenarios belonging to the same emotion category. The obtained correlation coefficients were 0.68 for deck A and 0.70 for deck B. In addition, we compared the average probability that a card was sorted together with the second one from the same emotion category, with the average probability that the two cards were sorted together with a card from a different category. The average within-category probabilities were 0.61 (sd = 0.20) for deck A, and 0.71 (sd = 0.17) for deck B, whereas the average between-category probabilities were only 0.05 (sd = 0.01) and 0.04 (sd = 0.02) for decks A and B respectively. Hence, a card was on average 11.5 (deck A) or 17.8 (deck B) times more likely grouped together with its partner from the same emotion category than with a card belonging to a different one. These results are comparable to the corresponding average within- and between-category similarities calculated for the data from study 2 (0.62 and 0.04, respectively, resulting in a same/different ratio of 15.5). Finally, the percentage of correct emotion labels (using all codable ones) was lower than that found in study 2 ($A_f = 0.54$ for deck A and 0.59 for deck B), but the modal label used to name the two scenarios from each emotion category was correct for all (deck B) or all but one (deck A) of the emotions. It may therefore be concluded that subjects' spontaneous discriminations between emotions on the basis of appraisal-relevant situational information were comparable to, or at least not substantially worse than those found in the first two studies.

GENERAL DISCUSSION

The aim of the studies reported in this article was to provide baseline data on people's ability to discriminate emotions from appraisal-relevant situational information. Taken together, the results of the studies suggest that subjects can discriminate rather well between emotions, at least those considered, on the basis of such information. Although the average obtained discrimination accuracy of about 65 to 70% is still considerably below perfect, the data suggest that this may be due to factors such as complexity of the eliciting situation (study 1) and disagreements on the meanings of emotion terms (study 2). This is supported by the finding that there were at least a few scenarios for nearly all emotions on which nearly perfect agreement was obtained (study 1).

As explained in the Introduction, these baseline data can serve as a standard of comparison for the results obtained in empirical studies of structural models of appraisals. It is evident that, even if one uses the average discrimination scores as the baseline, the empirical results obtained so far are still fairly below that baseline. How can this discrepancy be explained? In the Introduction, several possible reasons were mentioned: (1) some of the examined emotions may not be clearly distinguished

by people or there may be substantial inter-individual differences in the use of emotion labels; (2) the distinctions drawn by people between different emotions may in part be based on nonappraisal elements; (3) important dimensions of appraisal may be missing in current appraisal models; and (4) methodological problems could have deflated the associations between emotions and appraisals. As mentioned, some support for (1) was obtained in study 2, but in general, the role of this factor, as well as that of (2) seem to have been minor for the emotions considered here. However, these factors do remain a possibility for other emotions, particularly emotions that appear to be highly similar (e.g. Smolenaars & Schutzelaars, 1986/87). In the remainder of this discussion, however, we concentrate on factors (3) and (4), by briefly reviewing the evidence that exists to date on the potential relevance of these factors.

Additional Dimensions of Appraisal

Suggestive evidence that current appraisal models may omit important dimensions of appraisal (with different models omitting different ones) comes from several studies. Tesser (1990), in a replication of Smith and Ellsworth's (1985) study, found that the inclusion of additional potential appraisal dimensions (that are, however, regarded as "situational features" by Tesser), namely personal relevance, interpersonal distance (close vs. distant other), and social comparison (self outperforms other vs. other outperforms self; a dimension that could be reconstructed as a form of selfevaluation) improved the accuracy of prediction for several of the emotions studied. Frijda et al. (1989) also included several further appraisal dimensions in addition to those proposed by Smith and Ellsworth. Although direct comparisons with the Smith-Ellsworth dimensions are not reported, the fact that predictive accuracy (chance-corrected) was about as high as that obtained by Smith and Ellsworth, even though twice as many emotions were included, suggests that the additional dimensions significantly contributed to emotion prediction. Reisenzein and Hofmann (1990), using a modified version of Kelly's (1955) repertory grid technique, obtained evidence for the relevance of further potential appraisal dimensions not explicitly considered so far in most empirical studies, in particular focus (event affects primarily the self vs. other) and evaluation of other/quality of social relationship. They also found that the inclusion of these dimensions enhanced prediction of emotions as compared with a subset of dimensions also used by Frijda et al. (1989). In fact, the statistical classification accuracy obtained in this study (in which the same 23 emotions were included as in the present investigations) was 62% (chance-corrected) and hence approximated the average accuracy obtained in the present study 1. However, Reisenzein and Hofmann's study differed from others also in further respects (the events *typically* conducive to the emotions had to be judged, and an attempt was made to specify more precisely the objects of appraisals). In a recent investigation, Reisenzein and Spielhofer (submitted) had subjects rate remembered emotion-eliciting events for 30 emotions on 22 scales designed to tap potential appraisal dimensions; for the 22 emotions also included in the present studies, 56% (chancecorrected) correct classifications were obtained. An interpretative problem that arises in this context is that several of the further potential dimensions of appraisal included in these studies,

although improving emotion discrimination, may not really be dimensions of *appraisal* at all (cf. Lazarus & Smith, 1988; Reizenzein & Hofmann, 1990; Tesser, 1990). A decision on this issue presupposes that clear criteria existed for considering a proposed dimension, a dimension of appraisal (as opposed to, particularly, a dimension denoting "situational features"), in addition to discriminatory potential. Several suggestions for such criteria have been made (e.g. Lazarus & Smith, 1988; Reizenzein & Hofmann, 1990; Reizenzein & Spielhofer, submitted), but this issue is not yet settled. However, even if several of the further dimensions measured in these studies were ultimately excluded from the category of appraisal dimensions, their results do support the point of the present studies that an appreciable amount of variance in emotions due to situations is not accounted for by current appraisal models.

Methodological Factors

In the Introduction, three kinds of methodological factors were mentioned that could have deflated the relations between emotions and appraisals: (1) in most existing studies of emotional appraisals, subjects were asked to judge the total "emotional situation" on the appraisal scales; this could have blurred the associations obtained between appraisals and emotions; (2) the scales used to measure appraisals may have suffered from low reliability or validity; and (3) the statistical prediction models used may have been partly inadequate.

Some suggestive evidence does exist for the relevance of all of these factors. Concerning (1), Reizenzein and Hofmann (1990) obtained higher prediction rates in a study in which it was attempted to specify more precisely the objects of appraisals (see also, Roseman et al., 1990). However, because no control group receiving standard appraisal instructions was included, the importance of this factor could not be clearly determined. Concerning (2), several studies found that some of the scales used to measure appraisal dimensions had low internal consistencies (Roseman et al., 1990; Tesser, 1990) and reliability (Reizenzein & Spielhofer, submitted). Also, they may not always have been used as intended by the researchers. Reizenzein and Spielhofer (submitted) found, for example, that an "expert" coding of the 460 scenarios used in the present study pointed to a number of significant differences in scale use as compared with that of naive subjects (who judged remembered emotioneliciting scenarios) and that statistical discrimination accuracy was substantially higher (chance-corrected 71%) for the expert coded situations.

Finally, concerning (3), the appraisal dimensions may not combine additively in affecting emotion judgements (Tesser, 1990), or the scales may not be linearly related to emotions (e.g. Price, Barrell, & Barrell, 1985). The existence of nonlinear relations between appraisals and emotions could have additionally contributed to the higher discrimination scores obtained by Reizenzein and Hofmann (1990) and Reizenzein and Spielhofer (submitted), as these authors used binary variables as predictors (cf. Neter & Wasserman, 1974). However, the current data are only suggestive. Future research in this area should explore these possibilities in greater detail.

REFERENCES

- Allison, P.D. & Liker, J.K. (1982). Analyzing sequential categorical data on dyadic interaction: A comment on Gottman. *Psychological Bulletin*, *91*, 393-403.
- Averill, J.R. (1975). A semantic atlas of emotional concepts. *JSAS catalog of selected documents in psychology*, *5*, 330 (MS No. 421).
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46.
- Dalkvist, J. & Rollenhagen, C. (1989). *On the cognitive aspects of emotions: A review and a model*. Report No. 703 from the Department of Psychology, University of Stockholm.
- Davitz, J. (1969). *The language of emotion*. New York: Academic Press.
- Ekman, P. (Ed.) (1982). *Emotion in the human face*. Cambridge University Press.
- Ellsworth, P.C. & Smith, C.A. (1988). From appraisal to emotion: Differences among unpleasant feelings. *Motivation and Emotion*, *12*, 271-302.
- Fehr, B. & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, *113*, 464-486.
- Fiehler, R. (1990). *Emotion und Kommunikation* [Emotion and Communication]. Berlin: de Gruyter.
- Fillenbaum, S. & Rapoport, A. (1971). *Structures in the subjective lexicon*. New York: Academic Press.
- Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*, 378-382.
- Frijda, N.H. (1986). *The emotions*. Cambridge University Press.
- Frijda, N.H. (1987). Emotion, cognitive structure, and action tendency. *Cognition and Emotion*, *1*, 115-143.
- Frijda, N.H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*, 212-228.
- Gehm, T.L. & Scherer, K.R. (1988). Relating situation evaluation to emotion differentiation: Nonmetric analysis of cross-cultural questionnaire data. In K. Scherer (Ed.), *Facets of emotion: Recent research*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc, pp. 61-77.
- Hubert, L. (1977). Kappa revisited. *Psychological Bulletin*, *84*, 289-297.
- Izard, C.E. (1977). *Human emotions*. New York: Plenum.
- Kelly, G. (1955). *The psychology of interpersonal constructs*. New York: Norton.
- Laucken, U. (1989). *Denkformen der Psychologie* [Types of psychological thinking]. Bern: Huber.
- Lazarus, R.S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, *37*, 1019-1024.
- Lazarus, R.S. & Smith, C.A. (1988). Knowledge and appraisal in the cognition-emotion relationship. *Cognition and Emotion*, *2*, 281-300.
- Manstead, A.S.R. & Tetlock, P.E. (1989). Cognitive appraisals and emotional experience: Further evidence. *Cognition and Emotion*, *3*, 225-240.
- Marx, W. (1982). Das Wortfeld der Gefühlsbegriffe [The semantic field of emotion concepts]. *Zeitschrift für experimentelle und angewandte Psychologie*, *29*, 137-146.
- Miller, G.A. (1969). A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology*, *6*, 169-191.
- Neter, J. & Wasserman, W. (1974). *Applied linear statistical models*. Homewood, IL: Irwin.
- Ortony, A., Clore, G.L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge University Press.
- Osgood, C.E. (1966). Dimensionality of the semantic space for communication via facial expressions. *Scandinavian Journal of Psychology*, *7*, 1-30.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds), *Emotion: Theory, research, and experience*, Vol. 1. New York: Academic Press, pp. 3-33.
- Price, D., Barrell, J.E., & Barrell, J.J. (1985). A quantitative-experiential analysis of human emotions. *Motivation and Emotion*, *9*, 19-38.
- Reisenzein, R. & Hofmann, T. (1990). An investigation of dimensions of cognitive appraisal in emotion using the repertory grid technique. *Motivation and Emotion*, *14*, 1-26.
- Reisenzein, R. & Spielhofer, C. (Submitted). Subjectively salient dimensions of emotional appraisal.

- Roseman, I.J. (1984). Cognitive determinants of emotion: A structural theory. In Shaver, P. (Ed.), *Review of personality and social psychology*, Vol. 5. Beverly Hills: Sage, pp. 11-36.
- Roseman, I.J., Spindel, M.S., & Jose, P.E. (1990). Appraisals of emotion-eliciting events: Testing a theory of discrete emotions. *Journal of Personality and Social Psychology*, 59, 899-915.
- Rosenthal, R. (1982). Conducting judgment studies. In K.R. Scherer & P. Ekman (Eds), *Handbook of methods in nonverbal behavior research*. Cambridge University Press, pp. 287-361.
- Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.
- Scherer, K.R. (1984). On the nature and function of emotion: A component process approach. In Scherer, K.R. & Ekman, P. (Eds), *Approaches to emotion*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc, pp. 293-317.
- Schmidt-Atzert, L. (1981). *Emotionspsychologie* [The psychology of emotion]. Stuttgart: Kohhammer.
- 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.A. & Ellsworth, P.C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, 48, 813-838.
- Smith, C.A. & Ellsworth, P.C. (1987). Patterns of appraisal and emotion related to taking an exam. *Journal of Personality and Social Psychology*, 52, 475-488.
- Smolenaars, A.J. & Schutzelaars, A.J.H. (1986/87). On 'cognitive' semantics of emotion words: Solomon quasi-ecologically tested. *Journal of Semantics*, 5, 207-231.
- Solomon, R.C. (1976). *The passions*. Garden City, NY: Anchor Press/Doubleday.
- Tesser, A. (1990). Smith and Ellsworth's appraisal model of emotion: A replication, extension, and test. *Personality and Social Psychology Bulletin*, 16, 210-223.
- Tomkins, S.S. & McCarter, R. (1964). What and where are the primary affects? Some evidence for a theory. *Perceptual and Motor Skills*, 18, 119-158.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer.