

The Autonomic Differentiation of Emotions Revisited: Convergent and Discriminant Validation

GERHARD STEMMER
University of Freiburg

ABSTRACT

The convergent and discriminant validity of three models of physiological emotion specificity were compared. Forty-two female students served as subjects in a 2 (Context of emotional inductions: real-life, imagery) \times 3 (Emotion: fear, anger, control) + 1 (Happiness induced in real-life context) repeated measures design. The dependent measures included self-reports of emotion, Gottschalk-Gleser affect scores, back and forearm extensor EMG activity, body movements, heart period, respiration period, skin conductance, skin temperatures, pulse transit time, pulse volume amplitude, and blood volumes. Self-report data confirmed the generation of affective states in both contexts, as intended. Planned multivariate comparisons between physiological profiles established discriminant validity for fear and anger in the real-life context, whereas under imagery, emotion profiles were essentially equal. Convergent validity could not be substantiated. Implications for models of physiological specificity of emotion were discussed.

DESCRIPTORS: Emotion, Specificity, Physiological profiles, Discriminant and convergent validation, Fear, Anger, Happiness, Affective imagery, Emotion induction, Multivariate psychophysiology.

Recent research has suggested that there is considerable specificity in the physiological responses associated with diverse emotional states (Ekman, Levenson, & Friesen, 1983; Fridlund, Schwartz, & Fowler, 1984; Roberts & Weerts, 1982; Schwartz, Weinberger, & Singer, 1981). Ekman et al. (1983), for instance, studied six emotions in two different experimental contexts, during a posed facial expression task and during reliving past emotional experiences. The physiological emotion profiles were significantly different, but an Emotion \times Context interaction indicated that the contexts were a significant determinant of the actual emotion profiles. This focus on the hypothesis that emotional states are associated with specific physiological profiles is reminiscent of the arguments of Ax (1953), Funkenstein (1956), and J. Schachter (1957), and marks

a departure from the dominant view that emotions are characterized simply by general and undifferentiated physiological arousal. On the other hand, the focus on effects that the emotion eliciting context might have on actual physiological emotion profiles, could eventually lead to a revision of the concept of physiological emotion specificity.

This current shift back to the specificity position has occurred even though—or perhaps because—most of the recent studies departed in important ways from the classical investigations, in that they extended the definition and operationalization of emotion. First, new concepts in emotion theory led to new predictions concerning the set of physiological responses expected to reflect emotion specificity. Whereas most of the older investigations in the context of Jamesian body-reaction theory (James, 1884; for a survey of emotion research, cf. Leventhal & Tomarken, 1986) sampled physiological activity from a large array of autonomic and somatic measures, the more recent facial feedback hypothesis (Laird, 1984) or hemispheric lateralization research (Tucker, 1981)—with facial electromyography and electroencephalography, respectively—used different and more circumscribed sets of physiological variables. Second, whereas the older investigations predominantly used real-life emotion inductions (i.e., staging a task during which the involved subject probably will be affected emotion-

This research was supported by the Deutsche Forschungsgemeinschaft (SFB 115, Projekt C21) while the author was at the Universitätsklinik Hamburg-Eppendorf, FRG. I gratefully acknowledge the contributions to this study made by Burghard Andresen, Eckhard Irrgang, Wolfgang Spehr, and Eckart Thom. I am indebted to Paul Grossman for comments on an earlier draft of this paper, and I appreciate the help of Judith Whittaker in polishing up the English.

Address requests for reprints to: Gerhard Stemmer, Universität Freiburg, Forschungsgruppe Psychophysiologie, Belfortstr. 20, D-7800 Freiburg, FRG.

ally), investigators recently favored affective imagery of personal life episodes (Brewer, Doughtie, & Lubin, 1980) or of standard affect situations (Schwartz & Weinberger, 1980). This induction technique is easy to apply, it poses no ethical problems, and it is expected to have an interindividually comparable affective impact (Roberts & Weerts, 1982). Third, strategies of psychophysiological assessment (Stemmler & Fahrenberg, 1989) and statistical data analysis are gradually improving. Examples of this improvement include the multivariate analysis of physiological emotion profiles (Fridlund et al., 1984) and the disentangling of intensity and pattern effects in profile analysis (Stemmler, 1988).

In spite of an appropriate variable set, an effective induction technique, and a suitable statistical strategy, most of the previous studies have not explicitly defined or operationalized emotion specificity. With a growing interest in the role of the context of an emotion induction, however, clearly defined concepts of emotion specificity are needed; three of them will be introduced in this paper.

The aim of this study was to investigate the differentiation of emotions: (a) Employing a convergent and discriminant construct validation approach, (b) varying the eliciting context, (c) using autonomic and non-facial somatic measures in the tradition of body-reaction theory, (d) comparing physiological profiles and not primarily single variables, and (e) casting experimental hypotheses and results into the framework of specificity models. We also attempted to relate this study to the main body of psychophysiological emotion research by selecting emotions and contexts that have been most often employed in previous work: Fear and anger were elicited in a real-life and in an affective imagery context; for reasons described in the *Experimental Design* section, happiness was induced in a real-life context only.

Situational Response Specificity and Physiological Emotion Specificity

Two aspects of physiological activation processes, *intensity* and *pattern*, have played a prominent role in emotion theory (cf. the positions of James, 1884, and Cannon, 1929) and in accounts on the specificity of physiological responses (e.g., Cacioppo & Petty, 1983; Davidson, 1978; Fahrenberg, 1986; Lacey, 1967). Of particular relevance here is the notion of situational response specificity (Engel & Moos, 1967), which postulates that different stimuli activate different physiological response patterns, notwithstanding any differences in the intensity of diffuse physiological activity. The principle of situational specificity points to the fact that

the situational context of any emotion induction *per se*, irrespective of any emotional stimulus, has the potential to influence the actual physiological profile. To claim physiological emotion specificity would therefore require that the impact of emotional stimuli, as compared to that of contexts, on physiological profiles is "sufficiently large" and consistent. The three specificity models discussed below assume clearly different degrees of influences from emotional stimuli on physiological profiles; all of them, however, presume the existence of specific context profiles. It should be noted that by no means are these three models the only ones conceivable; they were selected because they can be related both to different theoretical positions in emotion theory and to the construct validation approach adopted here (see below).

When trying to discriminate between context and emotion, operational definitions of these concepts should be given. As used in this study, the context of an emotion induction is the experimental setting as defined by the physical characteristics of the situation including instructions, the subject's bodily position as requested by the experimenter (e.g., closed eyes or a reclined position), previous stimuli, and time elapsed. Excluded from this definition of context are emotional stimuli (which may or may not lead to an affective state) and all behavioral or intrapsychic responses of the subject. *Emotion* is indexed by a self-report of emotion or by an expert rating of overt responses the subject emits (e.g., expressive behavior, content of speech). Excluded from this definition of emotion are physiological responses, because this study asks whether specific physiological profiles also belong to the nomological network of the construct of emotion. Note that these definitions serve only the experimental aims of this study, they do not follow from the emotion theories mentioned in the next paragraphs.

Specificity Models and Emotion Theory

Among specificity models placing obviously different weight on the notion of physiological emotion specificity, are the following three:

First, *emotion nonspecificity* proposes that physiological responses are not an integral part of emotion and that specific physiological emotion profiles do not exist or at least do not have a functional role for the development or maintenance of emotions. This position seems to be implied in cognitive theories of emotion (Lazarus, Kanner, & Folkman, 1980; Mandler, 1980; S. Schachter, 1975). Certainly, it does not preclude the existence of specific context profiles, nor an intensity variation of diffuse activation with differentially intense affective states.

Second, what might be called *context-deviation specificity of emotion* states that the context has a marked influence on the physiological profile, but that an emotion specifically modifies the intensity and/or pattern of this profile. Thus, an emotion is credited with the difference between the profiles representing the context alone and the added emotional stimulus. This specificity model is implied, for example, by Leventhal's "Perceptual-Motor Processing Model of Emotion" (Leventhal, 1979), in which the "discrepancy of spontaneous and volitional or intentional expressive activity is the basis for emotion" (p. 39). In Lang's "Bio-Informational Theory of Emotional Imagery" (Lang, 1979, 1984), this concept is formulated explicitly. Emotions are viewed as dispositions to activate an associative network in memory. These networks contain stimulus, response, and meaning information. Stimulus and response aspects "to a considerable extent [determine] pattern and amplitude of physiological responses in emotion" (Lang, Levin, Miller, & Kozak, 1983, p. 304). The meaning information adds an emotion-specific contribution (Foa & Kozak, 1986).

Third, *absolute emotion specificity* claims that an emotion is necessarily accompanied by a particular physiological pattern which should be clearly discernible over and above any situational specificity in a profile or any intensity variations of diffuse activation. The James-Lange emotion theories

share this position as well as most of the investigations in the tradition of body-reaction theory. For example, Ax (1953), Funkenstein (1956), and J. Schachter (1957) postulated that fear and anger could be distinguished by the epinephrine-norepinephrine ratio. Another example is the specificity of an elevated diastolic blood pressure during anger as compared to fear or rest (Schwartz et al., 1981).

A Construct Validation Approach

With the multitrait-multimethod matrix, Campbell and Fiske (1959) introduced a systematic method for the validation of constructs. This method was based on the idea that variables—although using different measurement methods—should correlate substantially if they converged on the same construct, and that variables—although using the same measurement method—should not correlate if they were indicators of different constructs (or discriminated between them). A multitrait-multimethod matrix can be used here, too, in order to explicate the hypotheses of this investigation, the general features of its design and statistical contrasts, and the expectations under the three specificity models (see Table 1).¹

¹I am grateful to Gerald P. Ginsburg who made this suggestion.

Table 1

*Discriminant and convergent validation of autonomic emotion profiles within and between contexts:
Contrasts and expected results under three specificity models*

Emotions	Real-Life				Imagery	
	Fear	Anger	Happiness	Control	Fear	Anger
Real-Life:						
Anger	NS=*					
	CS#					
	AS#					
Happiness	NS=	NS=				
	CS#	CS#				
	AS#	AS#				
Control	NS=	NS=	NS=			
	CS#	CS#	CS#			
	AS#	AS#	AS#			
Imagery:						
Fear	NS#					
	CS#					
	AS=					
Anger		NS#			NS=	
		CS#			CS#	
		AS=			AS#	
Control			NS#	NS=	NS=	
			CS#	CS#	CS#	
			AS#	AS#	AS#	

*NS = model of emotion nonspecificity, CS = model of context-deviation specificity, AS = model of absolute emotion specificity. An equal sign (=) indicates equal profiles, an unequal sign (#) indicates profile differences.

The following hypotheses, which previously have been investigated often in isolation, constitute the construct validation approach. *Hypothesis of Discriminant Validity 1.* Within contexts, physiological profiles of emotion inductions do not differ from profiles of control periods. It is assumed that profiles from properly selected control periods reflect the influence of the context of the emotion induction. More specifically, it is hypothesized that physiological profiles of fear, anger, and happiness induced in a real-life context are equal to a control profile of the same context, and that physiological profiles of fear and anger under imagery are equal to a control imagery. A failure to reject this hypothesis would indicate that emotion inductions effected neither changes in the intensity nor in the pattern of activation, over and above the activation elicited by the context. This is what would be expected under nonspecificity, whereas under the absolute and the context-deviation specificity models, the null hypothesis should be rejected.

Hypothesis of Discriminant Validity 2. Within contexts, physiological profiles of emotion inductions are undifferentiated, although they may vary from a control profile. More specifically, the emotions of fear, anger, and happiness elicited in a real-life context have identical physiological profiles, which is similarly posited for fear and anger elicited under imagery. If this hypothesis cannot be rejected, then different emotion profiles are not specifically patterned; yet they could be characterized by an equal increase in the intensity of diffuse activation. This possibility could be tested with the previous hypothesis. Again, the null hypothesis is expected under nonspecificity and its rejection is expected under the other two specificity models.

Hypothesis of Convergent Validity. Between contexts, physiological profiles of inductions of presumably identical emotions are different. More specifically, the physiological profiles of fear (anger) elicited in a real-life and in an imagery context are different. Failure to reject the null hypothesis would indicate either a markedly different influence on the physiological profiles of the contexts (which can be tested with the following context hypothesis), or a lack of convergence because of differences in emotion profile intensities or patterns. This result would be expected under both nonspecificity and context-deviation specificity. If the null hypothesis is rejected, then either (a) the two contexts possess very similar profiles (and emotional stimuli have no marked influence on the profiles, which could be ruled out with the hypotheses of discriminant validity), or (b) convergent validity of the emotion profiles could be established, which is the result implied by absolute emotion specificity. It should be noted that this hypothesis is formulated in the

opposite direction to the usual formulation of a null hypothesis: this choice has consequences for the formal aspects of statistical testing (see the *Data Analysis* section).

Hypothesis of Context Specific Profiles. Physiological profiles of contexts are equal. Whether the two contexts possess situational specificity, is important for an appropriate interpretation of results from the convergent validity hypothesis. Rejection of the context null hypothesis is not, however, a logical prerequisite for the construct validation approach, because different contexts (in the definition of experimental settings given above) and not different context profiles are assumed.

The interrelatedness of these hypotheses suggests that the status of physiological emotion specificity has to be evaluated on the basis of the particular pattern of results displayed in the multitrait-multimethod matrix. The patterns implied by the three specificity models are included in Table 1.

Method

Subjects

Fifty female medical students voluntarily participated in the study and were paid 80 DM. Eight had to be excluded because of equipment failure or prior information, leaving 42 subjects with a mean age of 23 ($SD = 2.5$) years. Subjects were recruited by notices describing the study as an investigation in "the physiological effects of talking, thinking, and listening" without any reference to emotion. The sampling of subjects was restricted to one gender, because gender effects have been observed in emotion research (e.g., Frodi, 1978). Female subjects were selected because studies on imagery vividness (e.g., Schwartz, Brown, & Ahern, 1980) suggested greater emotional self-report and/or physiological changes in women.

Experimental Design

The design included two within-subjects variables, Context (real-life, imagery) and Emotion (fear, anger, happiness, and control). The variables were not completely crossed (there was no happiness imagery: Because of another purpose of the study each real-life emotion induction had to be paired with each of the imagery tasks).² This requirement restricted the number of recollections of private emotive episodes to a maximum of three (i.e., the fear, anger, and control imagery), because four recollections in one session were expected to be too strenuous for the subjects.

The order of affective imagery was completely permuted; real-life inductions, however, were always presented in the order fear, anger, happiness. The congruence of emotion induction and the common feel-

²The study also investigated the interactive effects on rated affects in Gottschalk-Gleser speech samples of emotion induction and the subsequent instruction to recall private emotive episodes (Stemmler, Bruhn, & Koch, 1986; Stemmler, Thom, & Koch, 1986).

ings of uncertainty and fear at the beginning of an experiment and of relief at the end was expected to increase the probability of a successful induction. Inducing happiness first and fear last is psychologically counterintuitive.

In order to reduce potential order effects for the real-life inductions, (a) the first experimental phase (10-min duration including the scoring of self-reports), expected to include most of the experiment's "initial effects," was not used in the statistical analysis, and (b) trends in the physiological data were mostly excluded. As a check on remaining order effects, the stability of the physiological profiles of a repeatedly presented vigilance task was determined.

Setting and Apparatus

The experimental room was a soundproof, air-conditioned chamber with a largely nontechnical appearance. Subjects sat comfortably in a reclined position. Electrodes and transducers were connected to a box (Toennies) behind the seat, where the electrophysiological signals were preamplified with a gain of 100 (input impedance 100 MΩ). Impedance pneumography was measured with a Hellige Servomed system. Other equipment included Revox loudspeakers and a screen for slide projection, both in front of the subject. In an adjacent room were placed Toennies AC and DC amplifiers, couplers, audiovisual equipment, a 20-channel Medelec Van Gogh polygraph (System 50000), and ZAK timers and interfaces. A Data General Eclipse S230 computer with 12-bit AD conversion was used to store data on digital magnetic tape and to control the experiment.

With the exception of interviews, subjects stayed alone in the chamber. Whenever necessary, experimenter and subject communicated via intercom.

Procedure

Session I. Subjects completed a questionnaire on demographic variables, the Freiburger Persönlichkeitssinventar (Personality Inventory; Fahrenberg, Selg, & Hampel, 1978), and the Freiburger Beschwerdenliste (Complaint List; Fahrenberg, 1975). Subjects described themselves as less aggressive and mentioned more bodily complaints than the respective norm samples. Because the subject sample did not possess deviant scores on neuroticism and extraversion, selection effects were estimated to be tolerable.

Session II. Two to 12 weeks later, subjects became acquainted with the experimental room, the physiological recording devices, and two tasks (speech and number task) that were external to the emotion induction. This dry-run session was intended to reduce primacy effects at the beginning of the main experiment.

Session III. Four weeks after session II, subjects performed the main experiment (one subject per day from 9 a.m. to 1 p.m.). Upon arrival, subjects completed a questionnaire about sleep the previous night, medication, the consumption of alcohol, nicotine, and coffee, physical and psychological well-being that day, and other information (cf. Stemmler, 1984). (Approximately 10% of the subjects described themselves as

not particularly fit that day.) Then electrodes and transducers were positioned.

The experiment consisted of seven phases, each comprising several experimental conditions. Within each phase, physiological data were continuously recorded. Between phases, retrospective self-ratings of emotion were obtained. After the second, fourth, and sixth phases, the experimenter conducted a semistructured interview about the preceding emotion inductions. A final interview terminated the experiment.

Phase 1 (Familiarization with the experimental setting). After a 30-s prestimulus period, subjects waited for the completion of a technical check (2 min). Then the number task already known from Session II was again explained and performed. Two target digits that were continuously projected onto the screen had to be matched with auditorily delivered five-digit numbers. If the target digits were contained in the number, subjects had to press a miniswitch. Twenty numbers were presented within 90 s and were followed by a 30-s poststimulus period. Previous experience from a pilot study with similar subjects had shown that this task demanded a medium level of vigilance and after little training was not affectively demanding.

Phase 2 (Real-life induction of fear, speech task, affective imagery). After the prestimulus period, the following instruction was given: "Right now you will hear a story, because something will happen to you that is out of your control." This sentence contains elements often considered to be constituents of fear (Epstein, 1972; McReynolds, 1976; Seligman, 1975): Ambiguity of meaning, threat of something unknown, lack of control. Then the lights in the experimental room were dimmed and a self-produced "radio play" was started (8 min). It contained the last section of Edgar Allan Poe's *The Fall of the House of Usher* dramatically recited by an actor. The story expresses feelings of terror, morbid anxiety, and death. The recitation was mixed with part of Prokofiev's *2nd Symphony*, music of nervous and dramatic character. The fear induction ended with an unexpected darkness (1 min) in complete silence, followed by a 1-min poststimulus period with lights turned on.

The speech task was the introductory part of affective imagery. According to the standard procedure of Gottschalk and Gleser's (1969) content analysis of speech samples, subjects were asked to speak about personal life experiences for 5 min. For subsequent fear imagery, subjects were requested to recollect and speak about a frightening life-episode; for subsequent anger imagery, it was an annoying life-episode; for the control imagery, the standard instruction was used, "please talk about something interesting and exciting." For later transcription and affect rating, speech samples were tape-recorded. At the end of the speech period, subjects were asked to imagine their story vividly, with eyes closed (1 min).

Phases 3, 5, and 7 were identically composed of a 30-s prestimulus period, a number task (each one with different target digits and numbers), and a 30-s poststimulus period.

Phase 4 (Real-life induction of anger, speech task, affective imagery). The anger induction was embedded

in a task of solving five-letter anagrams, for example "ROWNB" (BROWN). After explaining the task, the instruction said:

Please pronounce the solution as soon as you find it. If you do not find it, please say, "I don't know." Each word will be screened for only five seconds. The task is not that simple, it is often found in intelligence tests. Remember to pronounce the solution loudly or to say, "I don't know." The next word will not appear until we have heard your answer.

Anagrams were displayed for 5 s, and the first 15 items were solvable (taken from the Leistungsprüfungssystem; Horn, 1962). Then, unexpectedly, the task was interrupted, "Listen, we can hardly understand you, although the amplifier's volume is already turned up to maximum. It would be best if you speak louder!" The following anagram was solvable. Then, in a brusque voice, a second interruption, "Louder, please!" The next anagram was unsolvable (taken from Boucsein & Frye, 1974). After the subject's answer, it was aggressively insisted, "Can't you speak up?" The last eight items were solvable and unsolvable mixed, followed by two 1-min poststimulus periods. The induction's psychological elements were supposed to be (a) frustration and shame (inability to solve some items of an intelligence test and to announce them loudly), (b) the loud speaking (expressive element of anger; Gellhorn, 1970), and (c) the experimenter's aggressiveness generating counter-aggression (Epstein & Taylor, 1967). Phase 4 ended with the speech and affective imagery tasks.

Phase 6 (Real-life induction of happiness, speech task, affective imagery). The subject was told that the physiological recordings thus far were free from errors and that therefore none of the experimental conditions had to be repeated, as sometimes purportedly had happened. The experimenter inquired concerning the subject's well-being, and then he announced a short rest, the forthcoming end of the experiment, and an extra monetary bonus (10 DM), followed by two 1-min post-stimulus periods. Then the speech and affective imagery tasks ensued.

Operationalizations. Physiological profiles for real-life emotions were taken from the initial 1-min post-stimulus periods directly following each of the emotion inductions. These periods were characterized by an identical momentary experimental setting, in that subjects sat quietly and simply waited. The same was true for the 30-s poststimulus periods following the number tasks, which were chosen as the control period for the real-life context. Profiles for affective imagery were taken from the 1-min imagery periods, and the imagery control period was the imagery following the speech task's standard instruction. Profile stability was checked across the number task periods, and their poststimulus periods were also used for trend reduction.

Self-Report Variables

Self-Report of Emotion. Because self-reports of emotion had to be completed 17 times, a short 12-item list (with a 7-point intensity rating) was administered.

Items were: (1) "I clearly remember the experimental condition," (2) "I clearly remember my emotions," (3) "I was happy," (4) "I was angry about someone or something," (5) "I felt fear," (6) "I was angry at myself," (7) "I was irritated," (8) "I was exhausted," (9) "I sensed my heart, stomach, or muscles," (10) "I was depressed," (11) "I was excited," (12) "The period was pleasant." These items were thought to represent the emotions of interest. A factor analysis (principal components analysis, followed by oblique direct oblimin rotation; assignment of items to factors according to largest loading of an item; all item communalities exceeded 0.50) demonstrated that Fear (items 5, 9, 10, 11), Anger (items 4, 6, 7), and Happiness (item 3, 12) could indeed establish independent dimensions, besides Exhaustion (items 8, 11)³ and a Control-Items factor (items 1, 2). For further analysis, the marker items of the first four factors (items 3, 4, 5, and 8) were selected. Self-reports of emotion were obtained for the system-check, the four number tasks, the three real-life emotion poststimulus periods, the three speech tasks (separately for before and during the talks), and the three affective imageries.

Short Statement About Affect. Before completing the affect list, subjects were asked to briefly describe their feelings. The responses were recorded and categorized according to emotion relevant key words as the subjects used them (i.e., the coder did not have to infer the meaning of the recordings, cf. Stemmler, 1984).

Gottschalk-Gleser Affect Scores. Content analysis of the speech tasks' verbal samples was performed by two experts who in previous studies (Koch & Schöfer, 1986) achieved an average interrater reliability of .85. Of the 18 scales, the two summary scales "Total Anxiety" and "Total Aggressivity" plus the "Hope" scale of positive affects were chosen for the present analysis.

Interview. The interviews following phases 2, 4, and 6 explored, with standard questions but in a free response format, experiential aspects of the real-life emotion inductions (e.g., "What did you experience during darkness?"). The response categories were similar to those for the affect statements.

Physiological Variables

Only some of the physiological variables recorded were suited for the comparison of autonomic emotion profiles within and between contexts. First, the two contexts involved eyes-open (real-life) and eyes-closed (imagery) conditions which prohibited the use of physiological variables sensitive to this *between-context* variation. This concerned the electroencephalogram (EEG), the electro-oculogram, the pulse volume amplitude from photoplethysmography at the forehead (because of frequent movement artifacts during eyes-open conditions), and the electromyogram (EMG) from frontalis and orbicularis oculi sites (cf. Grossberg & Wilson, 1968). Second, for an evaluation of autonomic emotion specificity, physiological measures strongly related to other theoretical backgrounds, as is facial EMG to the facial feedback hypothesis, should

³Item 11 loaded negatively on the Exhaustion factor.

be excluded. Otherwise different output systems of emotion, such as facial expression and autonomic reactivity, would be confounded in the comparison of emotion profiles. This consideration argued against the use of EEG and facial EMG for *within-context* emotion comparisons.

EMG was recorded from the extensor digitorum and the trapezius muscles (after Lippold, 1967). Whereas Ekman et al. (1983) has not found differences among six emotions in the forearm flexor EMG, Chessick, Bassan, and Shattan (1966) have reported greater neck muscle activity during anger than during fear. Fear as compared to a control period has been characterized by higher EMG extensor and trapezius activity (Oken et al., 1966).

Skin sites were carefully cleansed with alcohol. Ag/AgCl electrodes, with a surface of 0.5 cm^2 , were filled with Hellige electrode jelly. Filters were set to 0.53–3000 Hz, amplification was 2×10^4 . EMG couplers performed low frequency (below 50 Hz) filtering to remove slow artifacts, rectification, and smoothing by a Bessel low-pass filter (20 Hz). After 16 Hz digitization values were summed to give 0.5-s scores with the dimension microvolt.

Body movements, as gross indicators of somatic activity, were recorded with two accelerometers (Entrance Devices, Type No. EGA-125-100D), fixed with adhesive tape at the tip of the middle finger on the non-dominant hand and at the forehead. Amplification was about 10^5 , filters were set to 1.3–5.3 Hz, and the sampling rate was 64 Hz. Digital high-pass filtering (6-point Butterworth, eliminating frequencies below 2 Hz) and rectification yielded measures of acceleration in gravitational units.

Heart rate has a long history as a putative indicator of emotion. Compared to rest, fear has been reported to elicit higher heart rates (Chessick et al., 1966; Grossberg & Wilson, 1968; Oken et al., 1966; Tourangeau & Ellsworth, 1979), as has anger (Chessick et al., 1966; Frodi, 1978). During real-life emotion inductions, larger heart rate responses during fear than anger have been observed (Ax, 1953; Dykman, Ackerman, Galbrecht, & Reese, 1963; J. Schachter, 1957) as well as the converse (Chessick et al., 1966; Schwartz et al., 1981, during exercise). During affective imagery, heart rate reactivity often has not discriminated between fear and anger (Ekman et al., 1983; Roberts & Weerts, 1982; Schwartz et al., 1981).

Ag/AgCl electrodes were positioned right and left at the 11th intercostal space about 6 cm lateral to a vertical line through the mamillae. This lead yields a pronounced R-wave and allows for the simultaneous measurement of respiratory rate. Filters were set to 1.6–80 Hz, amplification was at a gain of 500, and the sampling rate was 256 Hz. Self-optimizing digital filters and dynamic criteria were used to obtain a high signal-to-noise ratio for R-wave recognition (Stemmler & Thom, 1979). Output was interbeat intervals in milliseconds.

Respiration rate discriminated in Ax's (1953) study between fear and anger (greater rate for fear); however, Chessick et al. (1966) and J. Schachter (1957) found no differences. In the present study, respiration was

recorded using impedance measurement via the EKG electrodes. Sampling rate was 16 Hz. Automatic parameterization yielded respiratory cycle times in milliseconds.

Skin temperature measurements at finger or forehead were not found in earlier studies to differentiate between rest, fear, and anger (Ax, 1953; Chessick et al., 1966; J. Schachter, 1957), but recently Ekman et al. (1983) has reported greater finger temperature changes in anger than fear. Here skin temperatures were recorded at the volar surface of the distal phalanx of the index finger on the non-dominant hand and at the forehead. The transducers and corresponding couplers were self-built with the aims of high resolution and precision, short response latency, linearity, and a negligible microclimate at the application site. The latter was achieved through the use of an open casing, which when fixed, slightly pressed the thermo-element onto the skin surface.

Peripheral pulses are interesting as indices of peripheral blood flow. For example, Smith, Houston, and Zurawski (1984) have recommended finger pulse volume as a measure of anxiety in response to threat of physical harm and social-evaluative threat. Here reflection-photoplethysmography was used to measure peripheral pulses at the volar surface of the distal phalanx of the middle finger on the non-dominant hand and at the left corner of the forehead. Transducers and couplers were self-built; characteristics are an optimized linearity, a temperature compensation circuitry, and an application reducing local circulatory effects (Irrgang, 1981). A low-pass filter (0.75 Hz) separated the pulse volume amplitude from the low-frequency blood volume component. Pulse volume amplitude (blood volume in parentheses) amplification was about 50 (5) and the sampling rate was 256 Hz (16 Hz). Pulse volume amplitude was automatically determined as the systolic peak's amplitude above the line connecting the neighboring diastolic points. Measures were in arbitrary units. Because of frequent movement artifacts, pulse volume amplitude from the head was excluded from further analyses.

Pulse transit time (PTT) was grossly estimated by the time (in milliseconds) between R-wave and the systolic peak of the finger pulse.

Electrodermal activity is often regarded to be one of the most prominent physiological variables to assess emotion, especially fear and anxiety. Comparing fear with anger, greater skin conductance reactivity (SCR) during fear (Ax, 1953), during anger (Dykman et al., 1963), or no differences (Chessick et al., 1966) have been reported. For skin conductance level (SCL), however, greater reactivity has been more often seen in fear (Ax, 1953; Dykman et al., 1963; J. Schachter, 1957) than in anger; but lack of discrimination has also been reported (Chessick et al., 1966).

Electrodermal activity was recorded at the volar surfaces of the proximal phalanges of the middle and the index fingers on the non-dominant hand. The measurement principle was constant voltage of 0.5 V with Ag/AgCl electrodes (surface 0.5 cm^2). The electrode site was rubbed slightly with isotonic Hellige jelly, the electrodes were fixed with tissue glue, and the

cup was filled with jelly through a small lateral hole, which guaranteed an exact effective electrode size. The coupler separated the DC component with a 10-s time constant and amplified it by about 10^4 (AC: 10%). The sampling rates were 16 Hz. Skin conductance responses with amplitudes larger than $0.01 \mu\text{Siemens}/\text{cm}^2$ were automatically detected. Skin conductance level was defined as the average DC level per 0.5 s in $\mu\text{S}/\text{cm}^2$. Skin conductance response was defined as the number of fluctuations per second.

Scoring of Physiological Data

Several variables were screened for artifacts by visual inspection. Artifacts were either recorded as missing data, or deleted in the case of spikes superimposed on otherwise smooth signals (skin temperature, pulse volume amplitude, blood volume, skin conductance response); for the interbeat interval and respiration period data, intervals were summed or subdivided if too many or too few events (e.g., the R-peak) had been detected (Cheung, 1981).

Raw Scores and Missing Data. Raw scores were defined as the arithmetic mean of the physiological data within each experimental condition. If the data within conditions were missing, which occurred in 4% of all combinations of conditions and physiological variables, the following procedure was applied: First, each subject's array of raw scores in each physiological variable was centered (subtraction of the mean), thereby excluding individual differences in overall means from the data. Second, for the experimental condition in question, the mean of these deviation scores was calculated from those subjects who had no missing data in the respective variable. Finally, this deviation-score mean was converted into a raw-score estimate⁴ which was substituted for the missing value.

Trends were intra-individually controlled by a moving baseline that was "anchored" at the four post-stimulus periods following the number tasks. Because

the latter task was supposed to have a constant and largely nonemotional impact on the subjects, different activation levels across the four poststimulus periods should largely reflect factors like habituation, somatic inactivity, or ennui that—given the nonpermutation of real-life emotion inductions—needed control.

Reactivity was defined as follows. First, difference scores were determined by subtracting from the raw scores the current baseline, which was obtained through linear interpolation between the neighboring baseline "anchors." Second, individual differences in the overall mean were excluded by centering (separately for each variable) the individual arrays of difference scores. Third, all centered difference scores of a variable were subjected to a McCall normalizing transformation (which did not change the order of means). Normalization can provide for comparable units of measurement in different variables (for a discussion of measurement models in psychophysiology, cf. Stemmler, 1987). To achieve a consistent direction of activation across variables, the polarity of interbeat interval, pulse transit time, finger pulse volume amplitude, and finger temperature was changed. Finally, all variables were standardized ($\bar{X}=50$, $SD=10$).

Data Analysis

Self-report and Gottschalk-Gleser affect scales were analyzed for the emotions effect separately within real-life and imagery contexts. Effects of the emotion inductions were analyzed with the test on parallelism of emotion profiles over affect scales, which—if rejected—would mean that an Emotions \times Affect Scales interaction existed. A significant interaction was followed by repeated-measures ANOVAs (using multivariate analysis of variance) and finally by planned univariate contrasts among means, which used their specific error terms (cf. Vasey & Thayer, 1987).

Physiological emotion specificity was tested by pairwise planned comparisons (see Table 1) of emotion profiles (which are defined as the vectors of scores of standardized physiological variables from the emotion inductions' measurement periods) using Hotelling's T^2 for correlated samples (Harris, 1985), which—if significant—was followed by repeated-measures ANOVAs and univariate contrasts among means. The critical alpha level for these tests depended on the particular null hypothesis and the number of pairwise comparisons constituting a hypothesis. For tests of discriminant validity, a family-wise alpha level of .05 was selected, which was obtained by Bonferroni adjustments (with three comparisons in a hypothesis, $p=.017$; with two comparisons, $p=.025$). For tests of convergent validity and profile stability, the null hypotheses postulated differences among means, so the family-wise beta error rate had to be kept low, which could be indirectly achieved by an elevated alpha level. Individual comparisons were tested at an alpha level of $p=.10$, which with two pairwise comparisons yielded a family-wise alpha error of $p=.19$. In order to get a post hoc indication of the actual power of these tests, power calculations (Cohen, 1977) using the obtained F -values were performed.

⁴By definition of deviation scores d_{ijk} ,

$$\Sigma_k d_{ijk} = \Sigma_k (x_{ijk} - m_{ij}) = 0, \quad (1)$$

with x_{ijk} the raw score of subject i , variable j , in experimental condition k ($k=1, K$), and m_{ij} the raw score average across conditions. In the case of H ($h=1, H$) missing and L ($l=1, L$) non-missing raw scores ($K=H+L$), H deviation-score substitutes \hat{d}_{jh} are obtained as the mean of deviation scores from those subjects whose data for that jk -combination were not missing. Then again Equation (1) should apply,

$$\Sigma_h \hat{d}_{jh} + \Sigma_l (x_{ijl} - \hat{m}_{ij}) = \Sigma_h \hat{d}_{jh} + \Sigma_l x_{ijl} - L \hat{m}_{ij} = 0, \quad (2)$$

with \hat{m}_{ij} an estimate of the unknown m_{ij} . Solving Equation (2) for \hat{m}_{ij} ,

$$\hat{m}_{ij} = (\Sigma_h \hat{d}_{jh} + \Sigma_l x_{ijl}) / L, \quad (3)$$

enables the application of the missing data replacement equation:

$$\hat{x}_{ijk} = \hat{d}_{jh} + \hat{m}_{ij}. \quad (4)$$

Results

Physiological Levels and Menstrual Cycle

Menstrual cycle may alter physiological levels and reactivity. Because a within-subjects design was used in this study, cycle time effects would not bias tests of the experimental hypotheses, unless an interaction of cycle time and experimental conditions was expected. However, cycle time effects, if present, would increase the error variance. As a cursory test, the distribution of subjects over the menstrual cycle and, for a limited number of variables, differences in physiological levels were studied.

Subjects (except 4 with cycle times over 30 days or amenorrhea) were grouped into six equally spaced classes over cycle time and found to be evenly distributed across classes, $\chi^2(5, N=38)=0.21 (p>.05)$. Furthermore, none of the three physiological variables tested (skin conductance response, interbeat interval, and finger temperature, following Bell, Christie, & Venables, 1975) had significantly different raw-score means in cycle time classes. These findings do not indicate an appreciable influence of cycle time on later results.

Number Tasks: Characterization and Test for Physiological Profile Stability

As a check on the internal replicability of physiological profiles as obtained in this study and on remaining order effects after the trend reduction procedure, number tasks from phases 3, 5, and 7 were compared.

Number of Errors. Five of the 20 numbers presented in each task contained the target digits. In order of task presentation, 39, 36, and 36 subjects recognized all five target numbers; 3, 4, and 3 subjects had four hits; 0, 2, and 3 subjects had three or less hits. This shows that the number task was not overly difficult. A chi-square test on the tasks \times hits frequency table was not significant, $\chi^2(4, N=42)=3.16$, which suggests that vigilance showed no decrement over the course of the experiment.

Self-Report of Emotion. The numerically low means of the four selected self-report items (Table 2) supported the view that the three number tasks did not elicit any intense feelings. Repeated-measures analyses showed that task repetitions did not significantly differ in Anger and Happiness scores; but they differed in Fear and Exhaustion scores, $F(2/40)=4.46, 4.26 (p<.05)$, which in both cases could be attributed to the contrast between the second and the later two number tasks, $F(1/41)=8.97, 8.71 (p<.005)$. A further comparison contrasting the third and fourth tasks was not significant. Thus, Fear scores dropped from the second to the following number tasks (which might support

Table 2
Self-reports of emotion

Condition	Mean Values			
	Fear	Anger	Happiness	Exhaustion
Number Tasks				
Number Task 2	0.88	0.64	1.71	1.26
Number Task 3	0.36	0.38	1.38	1.81
Number Task 4	0.29	0.50	1.31	2.12
Real-life Emotion Induction				
Fear	2.17	0.79	1.09	1.24
Anger	0.88	2.76	0.93	1.52
Happiness	0.12	0.36	4.17	1.59
Recall After Instruction				
Fear Episode	2.33	1.43	0.74	1.21
Anger Episode	0.95	3.17	0.76	1.21
Control Episode	1.26	1.26	2.29	1.07
Affective Imagery				
Fear Episode	2.07	1.38	0.41	1.50
Anger Episode	0.48	3.38	0.83	1.43
Control Episode	0.69	0.95	2.38	1.38

Note. Scale values range from 0 (not applicable) to 6 (completely applicable).

the previous argument for placing the fear induction at the beginning) and Exhaustion scores rose.

Stability of Physiological Profiles. Physiological profiles during the three number tasks are shown in Figure 1. Neither the set of orthogonal contrasts comparing (1a) number tasks two vs. three and four combined and (1b) tasks two vs. three, nor the set comparing (2a) tasks two and three combined vs. four and (2b) tasks two vs. three were significant, $F(14/28)=1.59, 1.33, 1.06$, and 0.81 , respectively ($p>.10$). This result suggests sufficient profile stability over the course of the experiment. Power analysis indicated that the chance of detecting profile differences in these comparisons, if they existed, was reasonably large, viz. $0.93(0.97)$ and $0.87(0.94)$ for the .10 (family-wise .20) alpha level.

First Context: Real-Life Inductions

Self-Report of Emotion. The means of the items, Fear, Anger, and Happiness, were elevated for the respective emotion inductions and low otherwise (Table 2). The test on parallel profiles of the emotion inductions over the above items indicated their significant interaction, $F(4/38)=43.36 (p<.001)$. Repeated-measures analyses yielded for Fear, Anger, and Happiness scores significant differences among the three emotion inductions; $F(2/40)=27.81, 19.61$, and $74.17 (p<.001)$. Planned contrasts confirmed that—compared to the other emo-

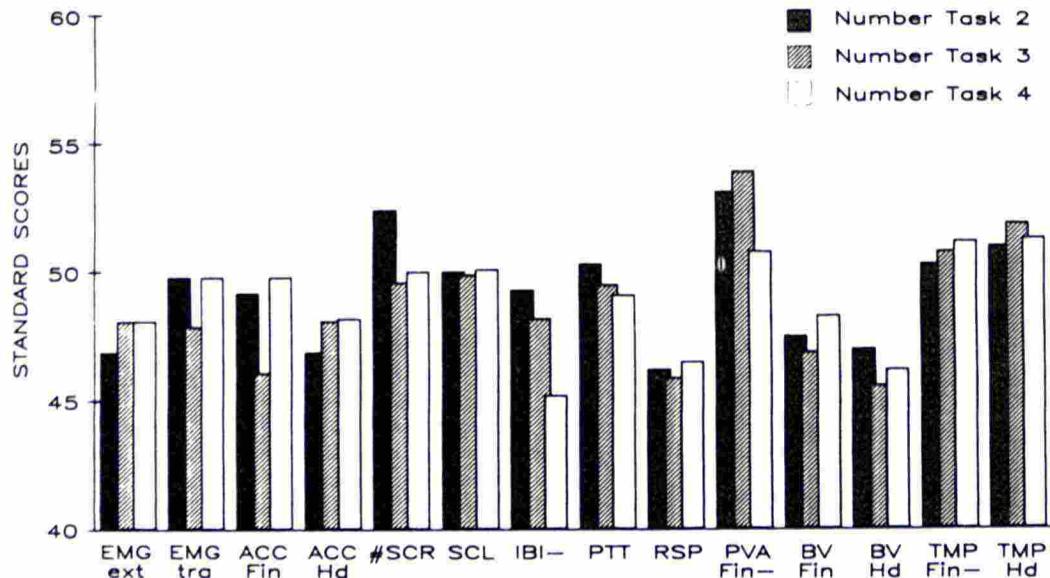


Figure 1. Physiological profiles during number tasks 2, 3, and 4. Values on the ordinate are normalized difference scores. Standardization ($\bar{X}=50$ and $SD=10$) is based on all subjects and experimental conditions. Variable abbreviations: EMG ext = EMG extensor digitorum, EMG tra = EMG trapezius, ACC Fin = finger acceleration, ACC Hd = head acceleration, #SCR = number of skin conductance responses, SCL = skin conductance level, IBI = interbeat interval, PTT = pulse transit time, RSP = respiration period, PVA Fin = finger pulse volume amplitude, BV Fin = finger blood volume, BV Hd = head blood volume, TMP Fin = finger temperature, TMP Hd = head temperature. A minus sign (-) indicates a change in the polarity of that variable.

tion inductions—the fear (anger, happiness) induction led to significantly larger means in the Fear (Anger, Happiness) scale; $F(1/41) = 24.87$ (38.98, and 148.71; $p < .001$). In contrast, Exhaustion scores did not significantly differentiate among the emotion inductions.

Short Statement about Affect. After fear induction, 55% of the statements were classified as fear- or threat-related, 12% expressed relief about the end of the fear induction, and 33% had no affective content. After anger induction, anger was expressed in 64% of the statements, 19% deplored an apparent inability to master the anagram task, and 17% said they were calm and relaxed. After happiness induction, 54% of the statements related to well-being and happiness, 20% revealed that the intention to induce emotions had been recognized, 14% expressed tiredness, and 12% had no affective content. Thus, the greater portion of the subjects spontaneously reported feelings as intended by the inductions.

Interview. Asked to recall their reactions to the various scenic elements of the emotion inductions, subjects could spontaneously report feelings, bodily reactions, or thoughts. During fear induction, the dramatic music elicited fear-related reactions in 53% of the subjects; the numbers for the other scenic elements were: paradoxical instruction 46%, darkness 41%, and the story itself 33%. For the an-

ger induction, 59% of the subjects reported aroused anger during the unexpected interruption ("Speak louder!"), 46% did so with regard to the apparatus "failure." Interestingly, 35% of the subjects explained their failure to solve some (truly unsolvable) anagrams with a distraction called forth by their anger. During happiness induction, the announcement of no repetitions was found enjoyable by 70% of the subjects, followed by the announcement of a rest (67%) and of a higher fee (57%).

Physiological Variables. Physiological profiles of the real-life emotion inductions and the control period are shown in Figure 2. Discriminant validity among the three emotion profiles could be established through the pairwise profile comparisons fear-anger, fear-happiness, and anger-happiness; $F(14/28) = 8.58$, 2.84, and 4.28, respectively ($p < .017$). Repeated-measures analyses showed univariate profile differences for EMG extensor, hand acceleration, skin conductance level, finger pulse volume amplitude, and finger and head temperature; $F(2/40) = 13.93$, 3.82, 5.26, 10.03, 4.48, and 25.32, respectively (all $p < .05$). Pairwise univariate contrasts indicated significant differences between fear and anger for EMG extensor, skin conductance level, finger pulse volume amplitude, and head temperature; $F(1/41) = 27.06$, 8.74, 19.76, and 42.03, respectively; between fear and happiness for head acceleration, skin conductance level, finger pulse

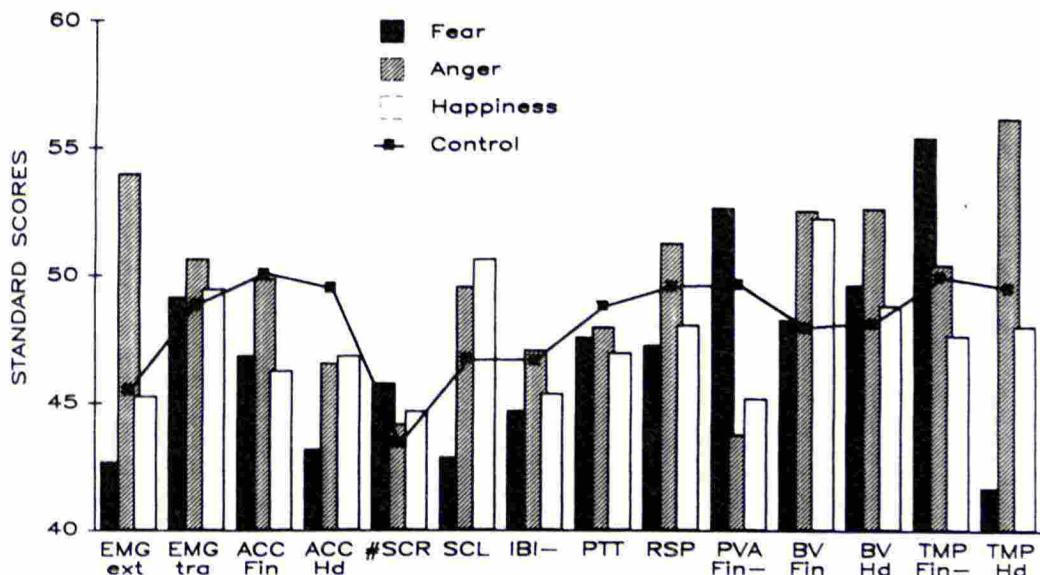


Figure 2. Physiological profiles during real-life emotion inductions of fear, anger, and happiness (bars), and during the real-life control condition (line). Values on the ordinate are normalized difference scores. See also legend of Figure 1.

volume amplitude, and finger and head temperature: $F(1/41) = 5.93, 9.95, 11.16, 9.18$, and 7.71 , respectively; and between anger and happiness for EMG extensor and head temperature: $F(1/41) = 16.44$ and 25.87 , respectively (all previous $p < .017$).

Discriminant validity between emotion and control profiles was significant for the comparisons fear-control and anger-control: $F(14/28) = 7.49$ and 7.37 ($p < .017$), but not for happiness-control. Pairwise univariate contrasts indicated significant differences between fear and control profiles for head acceleration, finger pulse volume amplitude, and finger and head temperature: $F(1/41) = 24.69, 6.10, 7.20$, and 13.09 , respectively; and between anger and control for EMG extensor, finger pulse volume amplitude, and head temperature: $F(1/41) = 21.65, 11.34$, and 12.95 , respectively (all previous $p < .017$).

Second Context: Imagery

Gottschalk-Gleser Affect Scores. The instructions to recall private emotive episodes of fear, anger, and "something exciting" generated, as intended, reports with these affective contents (see Table 3). The test on parallel profiles indicated a highly significant interaction of Gottschalk-Gleser scales and instructions, $F(4/38) = 32.32$ ($p < .001$). Subsequent repeated-measures analyses confirmed significant differences between instructions for the Total Anxiety, Total Aggressivity, and Hope scales: $F(2/40) = 33.86, 36.56$, and 15.70 , respectively ($p < .001$). Planned contrasts showed that—compared to the

Table 3
Gottschalk-Gleser affect scales

Topic of Recall	Mean Values		
	Total Anxiety	Total Aggressivity	Hope
Fear Episode	29.21	17.71	99.64
Anger Episode	20.93	28.36	99.38
Control Episode	16.12	15.71	100.12

other instructions—the fear (anger, control) instruction led to significantly larger means in Total Anxiety (Total Aggressivity, Hope): $F(1/41) = 69.00$ ($72.60, 26.62$; all $p < .001$).

Self-Report of Emotion During Talk. Another check on the differential affective impact of the private episodes was the self-report of emotion during recall (see Table 2). Instruction profiles over Fear, Anger, and Happiness scores were not parallel, $F(4/38) = 11.22$ ($p < .001$). Repeated-measures analyses found significant differences among instructions for Fear, Anger, and Happiness scores: $F(2/40) = 9.00, 14.25$, and 9.34 , respectively ($p < .001$), but not for Exhaustion scores. Planned contrasts confirmed that the instructions to generate fear and anger episodes led to large Fear and Anger scores, and that the control instruction generated emotionally positively toned recalls (large Happiness scores): $F(1/41) = 15.83, 28.02$, and 18.44 , respectively ($p < .001$).

Self-Report of Emotion During Imagery. Specificity of self-reported emotions could be established

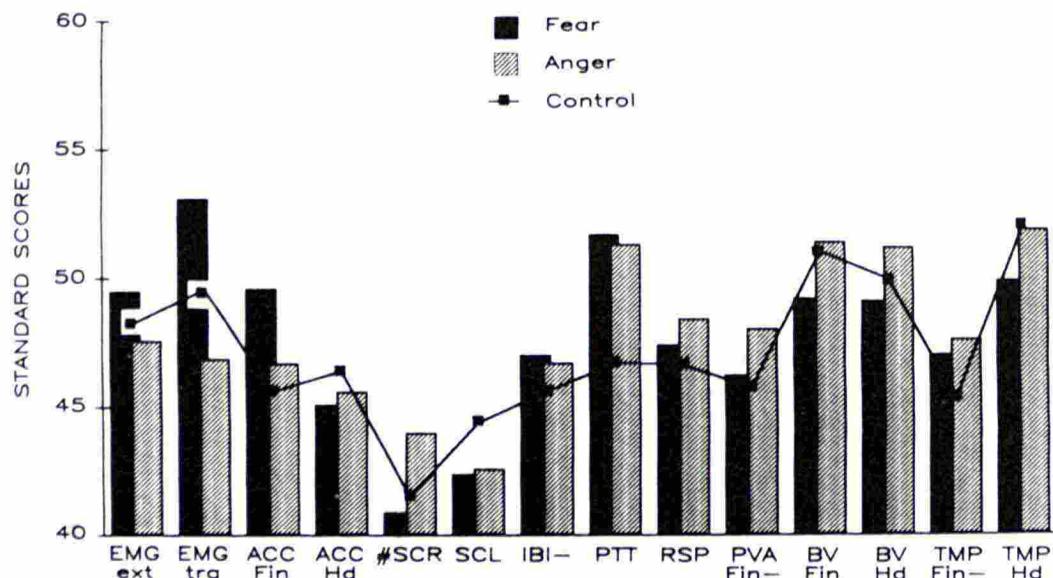


Figure 3. Physiological profiles during affective imagery of fear and anger episodes (bars), and during the imagery control condition (line). Values on the ordinate are normalized difference scores. See also legend of Figure 1.

also during imagery (see Table 2). Imagery profiles over Fear, Anger, and Happiness scores were not parallel, $F(4/38)=18.74$ ($p<.001$). Repeated-measures analyses yielded significant differences among imagery conditions for Fear, Anger, and Happiness scores: $F(2/40)=13.36$, 24.87, and 12.19, respectively ($p<.01$), but again not for Exhaustion scores. Planned contrasts also confirmed the specificity of affective imagery for Fear, Anger, and Happiness scores; $F(1/41)=24.88$, 50.09, and 21.12, respectively ($p<.001$).

Interview. Seventy percent of the subjects reported that during their speeches they were emotionally aroused; 59% ascribed to their speeches personal importance; 27% reported difficulty in talking about a personally relevant episode; 29% said they would often think about this episode; for 36% the reported episode had a larger affective impact on them than any other episode in their lives.

Physiological Variables. Physiological profiles during imagery periods are depicted in Figure 3. Discriminant validity could be demonstrated neither between fear and anger imagery profiles, $F(14/28)=1.39$, nor between affective imagery and control profiles, $F(14/28)=1.33$ and 1.09 for the fear-control and anger-control comparisons.

Comparison Between Contexts

Control Periods. Physiological profiles in real-life and imagery control periods were significantly different, $F(14/28)=2.67$ ($p<.05$).

Emotion Inductions. Convergent validity, as judged by the comparison of fear and anger profiles across real-life and imagery contexts, could not be

established; $F(14/28)=4.63$ and 3.27, respectively ($p<.01$, a probability that is far below the preset alpha level of .10). The actual power of these tests was above 0.999 (for the .10 and .20 levels).

Discussion

This investigation applied a convergent and discriminant validation approach to the study of autonomic and non-facial somatic emotion specificity, where the effects of different emotion inductions within and between contexts on self-reports of emotion and physiological profiles were determined. As a framework for this discussion, different specificity models have already been introduced. These models and their implications for emotion theory, experimental design, statistical analysis, and interpretation have not been provided in previous studies.

The specificity models are based on the notion that several factors may determine a particular physiological profile, the most prominent factor being *situational specificity*. Support for this idea was also found in our study in the different profiles of control periods within the two contexts. Situational specificity has been indicated to exert a consistent effect on physiological profiles (Lacey, 1962). Moreover, since a seemingly uniform condition like imagery has been said to split into several profile-influencing factors (Lang, 1984), the specific contributions of emotions, if any, have to stand against these other specificities. This experiment was designed to help disentangle these contributions.

Two obvious requirements for such a validation strategy were met. First, it could be shown that over

the course of the experiment, physiological profiles of a largely non-emotional task (the number task) were stable. This profile replicability substantiated the internal validity of the experiment. Second, without evidence of self-reported emotional specificity, the physiological data would have remained inconclusive. In this study, the self-report of emotion (e.g., Fridlund et al., 1984) and the experimenter's intent of inducing a particular emotion (e.g., Ax, 1953) jointly produced such evidence, providing two convergent sources of validity for the elicitation of specific emotional states. However, it might be noted that short statements of affect and interviews revealed that approximately one-third of the subjects did not explicitly express the intended feelings. As unfortunate as this may be, the aim of the present study was to analyze mean effects across subjects (cf. Fridlund et al., 1984, for an *intra-individual* approach). Nevertheless, the fact that the subjects were not complying to a higher degree with the experimental intentions might indicate that demand effects were at least tolerable.

Characteristic features of the real-life emotion inductions were: *Fear*—compared to anger, happiness, and control (exceptions noted)—was characterized by reduced head movements (except anger), low skin conductance levels (except control), marked finger vasoconstriction, low finger temperature (except anger), and low head temperature. *Anger*—compared to the other conditions—was characterized by tension in the extensor digitorum muscle, marked finger vasodilation (except happiness), and high head temperature. *Happiness* did not possess any discriminating features.

When the preceding summary of physiological profile characteristics is related to the literature, the impression of a bewildering inconsistency arises. In this study, heart rate, respiration rate, and number of skin conductance responses, which previously had been related to emotions, did not discriminate at all. During fear induction, we found larger skin

conductance levels than during anger, but the opposite was reported by Ax (1953). Contradictory to the epinephrine-norepinephrine hypothesis of fear and anger were the findings in peripheral vascular variables, showing finger vasoconstriction and temperature drops at the finger and head for fear but not for anger, which again is opposite to Ax's (1953) results. Ekman et al. (1983) also found colder fingers during fear induction, but unfortunately neither here nor in the Ekman et al. study has blood pressure been recorded. In sum, with such little agreement among different reported "emotion" profiles, any confidence in absolute autonomic emotion specificity is mitigated, if such inconsistencies are not systematically related to differences between studies in the precise quality of the aroused emotions. It should be recognized, for example, that "fear," "anger," or "happiness" actually are quite broad semantic categories.

The results of the *a priori* contrasts for physiological profiles are summarized in Table 4. The comparison with the expected results in Table 1 may help one to select the most adequate specificity model for these data. In the real-life context, physiological emotion and control profiles (except happiness vs. control) were significantly different. Together with the lack of convergence of profiles across contexts (as similarly seen in the significant Task × Emotion interaction in Ekman et al., 1983), this result then is in large accordance with the position of context-deviation specificity.

At first glance, this suggestion could be refuted with the argument that the "emotion" profiles, here taken from induction poststimulus periods, did not reflect an emotion-related process, but just the recovery from the behavioral demands of the preceding inductions. Indeed, a response diminution over the poststimulus periods could be demonstrated for the anger and, to a lesser extent, the happiness induction, but not for fear (Stemmler, 1984). However, the amplitude and duration of a recovery

Table 4
Discriminant and convergent validation of autonomic emotion profiles within and between contexts:
Results obtained with 14 autonomic variables

Emotions	Real-Life				Imagery	
	Fear	Anger	Happiness	Control	Fear	Anger
Real-Life						
Anger	#*					
Happiness	#	#				
Control	#	#		=		
Imagery						
Fear	#					
Anger		#				
Control				#	=	=

*An equal sign (=) indicates statistically equal profiles, an unequal sign (#) indicates statistically significant profile differences.

process might—but this remains to be demonstrated—also be influenced by an emotion. Such a modification of a context-dependent physiological profile by an emotion is exactly what the context-deviation hypothesis states.

Physiological profiles during affective and control imagery essentially were not discriminable. Together with the finding that self-reported emotions clearly showed specificity for the individualized affective imaginations and that convergence across contexts was not observed, this result is in accordance with physiological emotion nonspecificity. This result also would have been obtained if the imagery instructions had not led to actual emotions. But the self-reports, Gottschalk-Gleser scores, and interview data make this unlikely. Unfortunately there is no way to evaluate this assertion further, given the operational definition of emotion as used here.

The indiscriminability of physiological profiles during imagery could be explained with Lang's theory of affective imagery (Lang, 1979): It has been shown that response, rather than stimulus, propositions are effective in provoking physiological reactions. It could then be argued that in order to establish their imaginations, subjects might have used predominantly stimulus rather than response propositions, with resulting small physiological reactivity. In the present context, however, this argument would need the extension that under response propositions not just diffuse activation but specific physiological patterns were more likely to emerge. But this more crucial part of the argument did not find support in a study by Carroll, Marzillier, and Merian (1982) who noted "little specificity in proposition-response relationships" (p. 81).

What follows from these results for the concept of physiological emotion specificity? The overall pattern of results clearly stands against the position of absolute emotion specificity. In other words, a physiological profile alone does not seem to tell which, if any, emotion prevails. With regard to the fear and anger real-life inductions, the data are consistent with context-deviation specificity, whereas under imagery emotion nonspecificity seems to hold. Those who advocate some sort of emotion specificity might post hoc attribute more credibility to the real-life induction data and argue that the imagery was not intense enough (although this claim could not be substantiated with the self-report data, which show by and large equivalent means).

Those who advocate the nonspecificity position might primarily refer to the imagery data and propose that the context of the real-life inductions was only superficially identical and that the actual physiological profiles were differentially influenced by essentially non-emotional factors (but see the comments above concerning recovery processes in the light of context-deviation specificity).

However, these positions could be reconciled under the notion that—depending on emotion intensity—not only one but all of the specificity models might apply. With low intensity, emotions would have no particular influence on autonomic profiles and nonspecificity would prevail. With moderate intensity, emotions would interfere with the previous organism-environment interactions but would not redirect them. Then the situation-specific profile would be modified, as predicted under context-deviation specificity. With large intensity, emotions would disrupt the ongoing behavior (Mandler, 1980) and "install" another one, for instance, the "fight-flight" reaction. Here, emotions would largely influence autonomic profiles and absolute emotion specificity could hold. However, before this notion can be put to work, the meaning of "emotion intensity" must be clarified and operationalizations given.

A conceptual objection to the convergent-discriminant validation approach of this paper is conceivable. It could be postulated that a specific emotion is *intrinsically* tied to a particular context, rendering the independent experimental manipulation of context and emotion a questionable strategy. On the basis of such an alternative conceptualization, an approach different from the one adopted here must be specified.

This study demonstrated both the necessity and the utility of distinguishing among specificity models and related validation strategies. Elaborating further on the experimental context-emotion model as used in this paper, future research could contribute to a disentangling of the effects of context and emotion on physiological profiles by: (a) improving on the description and classification of physiological profiles of situations, that is, exploring the "map" of physiological situation specificity, (b) developing experimental paradigms with the aim of better context-emotion separation, and (c) exploring statistical models for the analysis of profiles and relating the aspects of intensity and pattern of profiles to characteristics of activation processes.

REFERENCES

- Ax, A.F. (1953). The physiological differentiation between fear and anger in humans. *Psychosomatic Medicine*, 15, 433-442.
 Bell, B., Christie, M.J., & Venables, P.H. (1975). Physiology of the menstrual cycle. In P.H. Venables & M.J. Christie (Eds.), *Research in psychophysiology* (pp. 181-207). London: Wiley.
 Boucsein, W., & Frye, M. (1974). Physiologische und psy-

- chische Wirkungen von Mißerfolgsstress unter Berücksichtigung des Merkmals Repression-Sensitization. *Zeitschrift für experimentelle und angewandte Psychologie*, 21, 339-366.
- Brewer, D., Doughtie, E. B., & Lubin, B. (1980). Induction of mood and mood shift. *Journal of Clinical Psychology*, 36, 215-226.
- Cacioppo, J.T., & Petty, R.E. (1983). Foundations of social psychology. In J.T. Cacioppo & R.E. Petty (Eds.), *Social psychophysiology: A sourcebook* (pp. 3-36). New York: Guilford.
- Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Cannon, W.B. (1929). *Bodily changes in pain, hunger, fear and rage*. New York: Appleton.
- Carroll, D., Marzillier, J.S., & Merian, S. (1982). Psychophysiological changes accompanying different types of arousing and relaxing imagery. *Psychophysiology*, 19, 75-82.
- Chessick, R.D., Bassan, M., & Shattan, S. (1966). A comparison of the effect of infused catecholamines and certain affect states. *American Journal of Psychiatry*, 123, 156-165.
- Cheung, M.N. (1981). Detection of and recovery from errors in cardiac interbeat intervals. *Psychophysiology*, 18, 341-346.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
- Davidson, R.J. (1978). Specificity and patterning in biobehavioral systems. *American Psychologist*, 33, 430-436.
- Dykman, R.A., Ackerman, P.T., Galbrecht, C.R., & Reese, W.G. (1963). Physiological reactivity to different stressors and methods of evaluation. *Psychosomatic Medicine*, 25, 37-59.
- Ekman, P., Levenson, R.W., & Friesen, W.V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221, 1208-1210.
- Engel, B.T., & Moos, R.H. (1967). The generality of specificity. *Archives of General Psychiatry*, 16, 574-581.
- Epstein, S. (1972). The nature of anxiety with emphasis upon its relationship to expectancy. In C.D. Spielberger (Ed.), *Anxiety, current trends in theory and research* (Vol. 2, pp. 291-337). New York: Academic Press.
- Epstein, S., & Taylor, S. P. (1967). Instigation to aggression as a function of degree of defeat and perceived aggressive intent of the opponent. *Journal of Personality*, 35, 265-289.
- Fahrenberg, J. (1975). Die Freiburger Beschwerdeliste FBL. *Zeitschrift für Klinische Psychologie*, 4, 79-100.
- Fahrenberg, J. (1986). Psychophysiological individuality: A pattern analytic approach to personality research and psychosomatic medicine. *Advances in Behaviour Research and Therapy*, 8, 43-100.
- Fahrenberg, J., Selg, H., & Hampel, R. (1978). *Das Freiburger Persönlichkeitseinventar FPI* (3rd ed.). Göttingen, FRG: Hogrefe.
- Foa, E.B., & Kozak, M.J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, 99, 20-35.
- Fridlund, A.J., Schwartz, G.E., & Fowler, D.C. (1984). Pattern recognition of self-reported emotional state from multiple-site facial EMG activity during affective imagery. *Psychophysiology*, 21, 622-637.
- Frodi, A. (1978). Experiential and physiological responses associated with anger and aggression in women and men. *Journal of Research in Personality*, 12, 335-349.
- Funkenstein, D.H. (1956). Nor-epinephrine-like and epinephrine-like substances in relation to human behavior. *Journal of Nervous and Mental Disease*, 124, 58-68.
- Gellhorn, E. (1970). The emotions and the ergotropic and trophotropic systems. *Psychologische Forschung*, 34, 48-94.
- Gottschalk, L.A., & Gleser, G.C. (1969). *The measurement of psychological states through the content analysis of verbal behavior*. Berkeley, CA: University of California Press.
- Grossberg, J.M., & Wilson, H.K. (1968). Physiological changes accompanying the visualization of fearful and neutral situations. *Journal of Personality and Social Psychology*, 10, 124-133.
- Harris, R.J. (1985). *A primer of multivariate statistics* (2nd ed.). New York: Academic Press.
- Horn, W. (1962). *Leistungsprüfsystem (LPS)*. Göttingen: Hogrefe.
- Irrgang, E. (1981). Untersuchungen zum Temperaturproblem physiologischer optischer Meßaufnehmer aus Halbleitern am Beispiel eines Reflektionsfotoplethysmografen. *Biomedizinische Technik*, 26, 106-109.
- James, W. (1884). What is emotion? *Mind*, 19, 188-205.
- Koch, U., & Schöfer, G. (1986). Testgütekriterien der Gottschalk-Gleser-Sprachinhaltsanalyse: Objektivität, Reliabilität und Validität im Rahmen deutschsprachiger Studien. In U. Koch & G. Schöfer (Eds.), *Sprachinhaltsanalyse in der psychosomatischen und psychiatrischen Forschung* (pp. 76-88). Weinheim: Beltz.
- Lacey, J.I. (1962). Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome. In E.A. Rubinstein & M.B. Parloff (Eds.), *Research in psychotherapy* (pp. 161-208). Washington, DC: American Psychological Association.
- Lacey, J.I. (1967). Somatic response patterning and stress: Some revisions of activation theory. In M.H. Appley & R. Trumbull (Eds.), *Psychological stress: Issues in research* (pp. 14-42). New York: Appleton-Century-Crofts.
- Laird, J.D. (1984). The real role of facial response in the experience of emotion: A reply to Tourangeau and Ellsworth, and others. *Journal of Personality and Social Psychology*, 47, 909-917.
- Lang, P.J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16, 495-512.
- Lang, P.J. (1984). Cognition in emotion: Concept and action. In C.E. Izard, J. Kagan, & R.B. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 192-226). Cambridge, MA: Cambridge University Press.
- Lang, P.J., Levin, D.N., Miller, G.A., & Kozak, M. (1983). Fear behavior, fear imagery, and the psychophysiology of emotion: The problem of affective response integration. *Journal of Abnormal Psychology*, 92, 276-306.

- Lazarus, R.S., Kanner, A., & Folkman, S. (1980). A cognitive-phenomenological analysis. In R. Plutchik & H. Kellerman (Eds.), *Emotion, theory, research, and experience* (Vol. 1, pp. 189-217). New York: Academic Press.
- Leventhal, H. (1979). A perceptual-motor processing model of emotion. In P. Pliner, K.R. Blankstein, & I.M. Spigel (Eds.), *Perception of emotion in self and others* (Vol. 5, pp. 1-46). New York: Plenum.
- Leventhal, H., & Tomarken, A.J. (1986). Emotion: Today's problems. *Annual Review of Psychology*, 37, 565-610.
- Lippold, O.C. (1967). Electromyography. In P.H. Venables & I. Martin (Eds.), *A manual of psychophysiological methods* (pp. 245-297). Amsterdam: North-Holland.
- Mandler, G. (1980). The generation of emotion: A psychological theory. In R. Plutchik & H. Kellerman (Eds.), *Emotion, theory, research, and experience* (Vol. 1, pp. 219-243). New York: Academic Press.
- McReynolds, P. (1976). Assimilation and anxiety. In M. Zuckerman & C.D. Spielberger (Eds.), *Emotions and anxiety* (pp. 35-86). Hillsdale, NJ: Lawrence Erlbaum.
- Oken, D., Heath, H., Shipman, W., Goldstein, I.B., Grinker, R.R., & Fisch, J. (1966). The specificity of response to stressful stimuli. *Archives of General Psychiatry*, 15, 624-634.
- Roberts, R.J., & Weerts, T.C. (1982). Cardiovascular responding during anger and fear imagery. *Psychological Reports*, 50, 219-230.
- Schachter, J. (1957). Pain, fear, and anger in hypertensives and normotensives. *Psychosomatic Medicine*, 19, 17-29.
- Schachter, S. (1975). Cognition and peripheralist-centralist controversies in motivation and emotion. In M.S. Gazzaniga & C. Blakemore (Eds.), *Handbook of psychobiology* (pp. 529-564). New York: Academic Press.
- Schwartz, G.E., Brown, S.-L., & Ahern, G.L. (1980). Facial muscle patterning and subjective experience during affective imagery: Sex differences. *Psychophysiology*, 17, 75-82.
- Schwartz, G.E., & Weinberger, D.A. (1980). Patterns of emotional responses to affective situations: Relations among happiness, sadness, anger, fear, depression, and anxiety. *Motivation and Emotion*, 4, 176-191.
- Schwartz, G.E., Weinberger, D.A., & Singer, S.A. (1981). Cardiovascular differentiation of happiness, sadness, anger, and fear following imagery and exercise. *Psychosomatic Medicine*, 43, 343-364.
- Seligman, M.E.P. (1975). *Helplessness. On depression, growth, and death*. San Francisco: Freeman.
- Smith, T.W., Houston, B.K., & Zurawski, R.M. (1984). Finger pulse volume as a measure of anxiety in response to evaluative threat. *Psychophysiology*, 21, 260-264.
- Stemmler, G. (1984). *Psychophysiologische Emotionsmuster*. Frankfurt: Lang.
- Stemmler, G. (1987). Implicit measurement models in methods for scoring physiological reactivity. *Journal of Psychophysiology*, 1, 113-125.
- Stemmler, G. (1988). Effects of profile elevation, scatter, and shape on discriminant analysis results. *Educational and Psychological Measurement*, 48, 853-871.
- Stemmler, G., Bruhn, M., & Koch, U. (1986). Eine psychophysiologische Validierungsstudie zum Gottschalk-Gleser-Versfahren: Einführung. In U. Koch & G. Schöfer (Eds.), *Sprachinhaltsanalyse in der psychosomatischen und psychiatrischen Forschung* (pp. 104-116). Weinheim: Beltz.
- Stemmler, G., & Fahrenberg, J. (1989). Psychophysiological assessment: Conceptual, psychometric, and statistical issues. In G. Turpin (Ed.), *Handbook of clinical psychophysiology: Vol. 1. Clinical psychophysiology: Theory and applications* (pp. 71-104). Chichester: Wiley.
- Stemmler, G., & Thom, E. (1979). Ein Programm für die interaktive Auswertung des Herzschlagabstands aus stark artefaktbehafteten EKG-Aufzeichnungen. *EDV in Medizin und Biologie*, 10, 122-127.
- Stemmler, G., Thom, E., & Koch, U. (1986). Zur Psychophysiologie von Gottschalk-Gleser-Sprachproben: Ein Beitrag zur Validierung des Gottschalk-Gleser-Versfahrens? In U. Koch & G. Schöfer (Eds.), *Sprachinhaltsanalyse in der psychosomatischen und psychiatrischen Forschung* (pp. 143-165). Weinheim: Beltz.
- 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.
- Tucker, D.M. (1981). Lateral brain function, emotion, and conceptualization. *Psychological Bulletin*, 89, 19-46.
- Vasey, M.W., & Thayer, J.F. (1987). The continuing problem of false positives in repeated measures ANOVA in psychophysiology: A multivariate solution. *Psychophysiology*, 24, 479-486.

(Manuscript received March 29, 1988; accepted for publication January 14, 1989)

Announcement

Behavioral Psychologist

The Department of Pediatrics, University of Tennessee, Memphis, seeks a Behavioral Psychologist with experience and publications in the field of cardiovascular reactivity and other cardiovascular risk factors. Position is full-time faculty with small practice portion negotiable. Experience with children preferred but not required. Level of appointment and salary commensurate with training and experience. Please direct all inquiries to: Bruce S. Alpert, M.D., Associate Professor and Chief of Cardiology Section, 848 Adams Avenue, Memphis, TN 38103. The University of Tennessee, Memphis is an Affirmative Action/Equal Opportunity Employer.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.