

BEYOND TEST ANXIETY: DEVELOPMENT AND VALIDATION OF THE TEST EMOTIONS QUESTIONNAIRE (TEQ)

REINHARD PEKRUN^{a,*}, THOMAS GOETZ^a, RAYMOND P. PERRY^b,
KLAUDIA KRAMER^c, MICHAELA HOCHSTADT^d and
STEFAN MOLFENTER^d

^a*Department of Psychology, University of Munich, Leopoldstrasse 13, 80802 Munich, Germany;*
^b*Department of Psychology, University of Manitoba, Canada;* ^c*Institute of Psychology, University
of Erlangen Nürnberg, Germany;* ^d*Institute of Psychology, University of Regensburg, Germany*

Measures of test emotions other than test anxiety are lacking. In a series of six studies, we developed a multi scale questionnaire assessing test related joy, hope, pride, relief, anger, anxiety, shame, and hopelessness (Test Emotions Questionnaire, TEQ). Consisting of subscales measuring affective, cognitive, physiological, and motivational emotion components, the scales can be used to identify both trait and state test emotions, and are available in German and English language versions. Using a rational empirical strategy of test construction, we first developed theoretical models of the component structures, antecedents, and effects of different test emotions. We then conducted two exploratory, qualitative studies on the occurrence and structures of these emotions. Finally, based on theory and our exploratory data, we constructed, analyzed, and revised the scales of the TEQ in four quantitative studies. Findings of correlational and confirmatory factor analysis indicate that the scales are reliable, structurally valid in terms of representing different test emotions and components within emotions, and externally valid in terms of correlating significantly with personality, learning, academic achievement, and perceived health problems. Many of these correlations proved to be stronger for test emotions other than anxiety, implying that test anxiety is neither the only relevant test emotion, nor necessarily the most important one.

Keywords: Test emotions; Test anxiety; Test Emotions Questionnaire; Learning; Academic achievement

Since the beginning of the last century, researchers have been interested in the emotional reactions of students taking tests and exams. During the early decades of the 20th century, it became clear that our civilization was about to become the civilization of testing it is today, making educational opportunities and life prospects dependent on individual success or failure on exams. American and Russian psychophysicologists (see Brown and van Gelder, 1938) as well as German psychoanalytical writers (e.g., Stengel, 1936) published articles on exam-related emotions. From these early beginnings, authors focused on the one emotion which seems to be most prominent in test-related subjective experiences, namely, anxiety.

Advances in research on test anxiety, however, had to wait for the development of suitable measurement instruments. Contrary to popular belief that test anxiety research

*Corresponding author. Tel.: +49 89 2180 5149; Fax: +49 89 2180 5250.
E mail: pekrun@edupsy.uni muenchen.de

originated in the 1950s (see Hembree, 1988), the first instrument measuring test anxiety was constructed in the 1930s by C. Brown at the University of Chicago (Brown, 1938). This instrument was a multi-item questionnaire measuring affective and physiological anxiety reactions to exams. Without referring to Brown's work, Mandler and Sarason (1952) later developed the Test Anxiety Questionnaire (TAQ), also measuring anxiety relating to tests and exams. The TAQ soon received widespread attention, and 1952 has since been regarded as the year in which test anxiety research was born. Research on the concept boomed for decades, and test anxiety became one of the most popular constructs in personality research. Test anxiety appears to have absorbed the energy of researchers to such an extent that test-related emotional reactions other than anxiety have received little more than cursory attention (see the findings of quantitative literature searches on achievement-related emotions in Pekrun and Frese, 1992; Pekrun *et al.*, 2002a).

Theoretically, however, tests and exams can be expected to trigger a wide variety of human emotions. Exam results are decisive for educational and occupational careers today, implying that high subjective values are attached to success and failure on exams. Since emotions are caused by events and objects bearing high positive or negative values, exams can be expected to trigger intense emotions. Exams are events which can be anticipated and recalled, implying that they can induce both prospective and retrospective emotions. The anticipation of an exam can create any range of subjective probabilities, as well as positive versus negative values of its expected course and outcome, thus being able to generate positive and negative emotions involving differing degrees of subjective certainty (such as hope versus anticipatory joy, and anxiety versus hopelessness). Furthermore, exams are frequently embedded in ongoing social interactions conveying social evaluations of individual achievement, implying that they can trigger not only individual emotions, but also social emotions, such as social comparison emotions involving appraisals of self-worth.

Beyond test anxiety research, the origins of test-related emotions have been analyzed by researchers exploring the appraisal antecedents of emotions. Specifically, emotions linked to exam outcomes have been used to test appraisal assumptions from three different theoretical paradigms: Weiner's (1985) attributional theory of achievement emotions, the reformulated helplessness and hopelessness theories of depression (Houston, 1995; Metalsky *et al.*, 1987), and transactional stress theory (Smith and Ellsworth, 1987). Studies on exam-related emotions conducted within these traditions have relied on the use of single rating scales or adjective lists to measure emotions. As in test anxiety research, however, any further progress in analyzing the structures, causes, effects, and remedies for test emotions other than anxiety may presuppose the development of instruments measuring these emotions and their components in more systematic ways.

The present research therefore aimed at developing a more comprehensive set of differentiated, multi-item self-report scales measuring a number of different test emotions (Test Emotions Questionnaire, TEQ). Test-related anxiety was to be taken into account along with other test-related emotions, including both negative emotions (anger, shame, and hopelessness) and positive emotions (joy, hope, pride, and relief). The scales of the TEQ were developed such that they can be used to measure both trait and state test emotions. In addition to the original German-language versions of the scales, English-language scale versions were developed as well. Scale construction was

accomplished through theory development and a series of six independent empirical studies based on the following strategy.

Strategy of Scale Construction and Overview of Studies

Any development of measures based only on theoretical considerations may be in danger of misrepresenting reality, and any development focusing on empirical data may be biased by implicit assumptions of test constructors and the random character of convenience samples of items and subjects. Therefore, we attempted to use both theory and empirical analysis to construct the TEQ scales. The rational-empirical strategy adopted consisted of the following steps.

Theoretical considerations. The development and validation of scales measuring a given construct presupposes a theoretical conception of the contents and boundaries of the construct, its internal structures, and its relations to external variables. In order to obtain such a theoretical foundation, we first attempted to define the term “test emotion.” We then conceptualized a classification of such emotions, and formulated assumptions about their component structures. Finally, we derived assumptions about the relations of test emotions to antecedents and outcome variables from Pekrun’s control-value theory of achievement emotions (Pekrun, 2000; Pekrun *et al.*, 2002a, b).

Exploratory analysis (Studies 1 and 2). To date, there is a lack of base rate knowledge on test emotions. We do not know which emotions are experienced in situations of tests and exams, how often they occur, or how they are structured (Zeidner, 1998). To gain empirical base rate knowledge for selecting emotions and constructing items, we conducted two exploratory studies analyzing occurrence and structures of test emotions.

Construction of the TEQ scales (Study 3). Based on our theoretical conception of the domain of test emotions and the findings of our exploratory studies, we constructed item pools for the first version of the TEQ. This version included scales for six major test emotions, namely, test-related joy, hope, relief, anger, anxiety, and hopelessness. These emotions were selected because of their relative frequency and their relevance for outcomes like achievement and health. In a sample of university students, we analyzed item and scale characteristics, also including external validity information on relations to demographic variables, personality traits, academic learning and achievement, and health complaints (Study 3).

Revision of the TEQ scales (Studies 4 and 5). Based on an analysis of item and scale characteristics, we revised the scales twice to produce a revised and then a final version of the TEQ scales (Studies 4 and 5). Again, university students served as samples. For each revision, item and scale analysis used score distributions, scale reliabilities, internal scale validity (convergent and divergent item validities, interrelations of scales), and external scale validity. Furthermore, two additional scales on test-related pride and shame were constructed and added to the revised TEQ because of the theoretical importance of these two emotions.

State versions, short versions, and English-language versions of the scales (Study 6). The TEQ scales can be used to measure either trait or state test emotions by adapting instructions accordingly. A more detailed description of the state version of the TEQ is

beyond the scope of this paper (see Molfenter, 1999). Also, we created short, eight- to 12-item versions of the scales. Finally, for international use and cross-cultural comparisons, the short versions of the TEQ scales were translated into English and tested in a sample of North American students (Study 6). Again, scale validation pertained to internal validity (interrelations of scales, confirmatory factor analysis of scale structures) as well as external validity.

THEORETICAL CONSIDERATIONS GUIDING SCALE DEVELOPMENT AND VALIDATION

Defining Test Emotions

Test anxiety is usually defined as anxiety subjectively relating to taking tests and exams, including anxiety related to the threat of failing an exam and the associated negative consequences (Zeidner, 1998). Conceptualized as an individual disposition to react to tests and exams in habitual ways through the experience of anxiety, test anxiety is regarded as a situation-specific personality trait (trait test anxiety; Spielberger *et al.*, 1976). As such, it can be seen as being part of generalized trait anxiety. As a momentary emotional state experienced before or while taking a specific test, it is a state emotion (state test anxiety).

These traditional definitions of trait and state test anxiety can be used to define test emotions more generally. Test emotions are defined here as *emotions subjectively relating to taking tests and exams*. Such emotions can be experienced at any time before, during, or after test taking. They are regarded as test emotions if they relate to the test, irrespective of when they are experienced. Predispositions to experience such emotions in habitual ways constitute *trait test emotions*. If emotions relating to a test or an exam are experienced at specific points of time before, during, or after that test, they can be seen as *state test emotions*.

Classifying Test Emotions

As for emotions more generally, *valence* and *activation* are two primary dimensions suitable to classify test-related emotions. Theoretically, these two dimensions are orthogonal, thus constituting a two-dimensional space (Feldman Barrett and Russell, 1998). Dividing both dimensions into two categories (positive versus negative, activating versus deactivating) creates four broad groups of test emotions: (a) *positive activating emotions* (e.g., enjoyment of the challenge implied by an exam, hope of success, pride in positive outcomes); (b) *positive deactivating emotions* (such as relief or pleasant relaxation after taking a test); (c) *negative activating emotions* (e.g., anger about the examiner, anxiety before an exam, or shame following failure); and (d) *negative deactivating emotions* (e.g., hopelessness when failure cannot be avoided). Valence and activation can be regarded as underlying dimensions which partially determine the effects of test emotions on achievement (see below).

A third important dimension is the *time reference* of an emotion. Emotions may relate in prospective (anticipatory) or retrospective ways to tests and exams. Major *prospective* test emotions are anticipatory joy, hope, anxiety, and hopelessness; major *retrospective* emotions may be outcome-related joy, pride, and shame. Both prospective

and retrospective emotions can be experienced at any time before, during, and after exams. This even applies to prospective emotions like anxiety (e.g., post-exam anxiety relating to imminent feedback on one's results) and retrospective emotions like relief (e.g., pre-exam relief relating to successful preparation). Typically, however, prospective emotions can be expected to peak before and at the start of the exam, and retrospective emotions after the exam. The TEQ seeks to measure test emotions from all four quadrants of the valence \times activation space, as well as both prospective and retrospective test emotions.

Components of Test Emotions

Emotions are commonly seen as systems of interrelated psychological and physiological component processes, important components being affective, cognitive, physiological, expressive, and motivational processes (e.g., Scherer, 1984). For example, the components of anxiety can be uneasy feelings of nervousness, worry cognitions, physiological activation, fearful facial expressions, and impulses to escape. Contemporary definitions of the test anxiety construct are in line with such a view. Since Liebert and Morris (1967) introduced the distinction between "worry" (cognitive) and "emotionality" (affective-physiological) components of test anxiety, researchers have adopted multi-component views of test anxiety.

Most often, it is just "worry" and "emotionality" which are distinguished as test anxiety components. However, from the perspective of general emotion research, this dichotomous conception is problematic for two reasons. First, it blurs the distinction between affective and physiological components of anxiety. This problem has successfully been addressed by I.G. Sarason (1984), who proposed to split the emotionality component into its "tension" and "bodily reactions" subcomponents, and presented a test anxiety inventory containing separate subscales measuring these subcomponents (Reactions to Tests questionnaire, RTT). Sarason's conception has guided recent research on the dimensionality of test anxiety (e.g., Hodapp and Benson, 1997).

Second, the two-component conception fails to take the motivational components of test anxiety into account. According to our everyday experience, thoughts and wishes to avoid failure and to escape exam situations are central to the experience of test anxiety. Therefore, early instruments measuring test anxiety did contain items assessing this component (e.g., Mandler and Sarason, 1952). In subsequent measurement development, however, these items were lost, and in contemporary measures such as the Reactions to Tests questionnaire (Sarason, 1984), the Test Anxiety Inventory (TAI; Spielberger, 1980), or the integrative scale devised by Hodapp and Benson (1997), the avoidance-related motivational component of test anxiety is no longer addressed. Traditional and recent conceptions of achievement motivation, however, take distinctions between approach and avoidance motivation into account (see Elliot and McGregor, 2001; Heckhausen, 1991), and address linkages between achievement motives and test anxiety (Elliot and McGregor, 1999). Test anxiety research might also profit from addressing the interface between anxiety and achievement motivation.

In line with components views of emotions, we propose that the *affective*, *cognitive*, *physiological*, and *motivational* components of test emotions be considered most important. To help item construction for the component subscales of the TEQ scales, we developed a theoretical taxonomy of these four components for the emotions to be

measured by the TEQ. This taxonomy was based on general accounts of the components of emotions (Scherer *et al.*, 1986), and on the findings of our exploratory analysis (Studies 1 and 2). Basically, the taxonomy consists of two axes: type of emotion and type of component. Affective, cognitive, physiological-expressive, and motivational components are subdivided in such ways that differences between test emotions can be specified.

For example, according to the component taxonomy, major *cognitive components* of test emotions include thoughts pertaining to (a) one's own competencies (including ability, knowledge, and preparation for the exam), (b) the demands of the exam (amount of material, difficulty, and time pressure), (c) expectations of performance on the exam, (d) expectations relating to the consequences of exam performance, (e) retrospective appraisals of performance during and after the exam, (f) the value and meaning of the exam material, and (g) the fairness of the exam's demands and procedures. Cognitive components of different test emotions are characterized by subsets of these dimensions and specific qualities of cognitions relating to them. For example, according to the taxonomy, whereas doubts about one's own competencies and expectancies of failure involving subjective uncertainty are typical of anxiety, expectancies of success involving uncertainty may be characteristic of hope, expectancies of success involving relative subjective certainty of joy, and certainty of failure of hopelessness (for a more detailed discussion of the taxonomy, see Molfenter, 1999).

Antecedents and Effects of Test Emotions: Assumptions of Pekrun's Control-Value Theory

Validity analysis of the TEQ was guided by the assumption that test emotions can exert profound effects on individual learning, achievement, personality development, and health. We used assumptions from Pekrun's control-value theory of achievement emotions (Pekrun, 2000; Pekrun *et al.*, 2002a, b) to analyze the functional construct validity of the TEQ scales with respect to learning and achievement. The theory implies that control- and value-related appraisals are major determinants of individual test emotions. An example is academic self-efficacy cognitions affecting emotions like exam-related joy, hope, anxiety, and hopelessness. Concerning effects on learning and achievement, the theory uses the two-dimensional classification of achievement-related emotions outlined above. Effects of emotions on motivation, strategies of learning and problem solving, the availability of cognitive resources (working memory capacity), and self-regulation are assumed to mediate their effects on performance and achievement.

According to the control-value theory, *positive activating* test emotions like enjoyment of challenge and hope of success can be expected to strengthen motivation, facilitate flexible strategies of learning and self-regulation, and preserve cognitive resources by helping the learner to focus attention on the task at hand, thus exerting beneficial effects on learning and exam performance. The *negative deactivating* emotion of hopelessness is assumed to be detrimental by reducing motivation, by only allowing shallow processing of task information, by consuming cognitive resources through worry and mental distraction from the task, and by undermining any strategic regulation of learning and problem solving, thus negatively affecting performance. For *positive deactivating* emotions like relief, as well as *negative activating* emotions

like anger, anxiety, and shame, more complex effects are assumed. For example, although test anxiety can reduce intrinsic motivation, flexible strategy use, and attentional resources, it may also prompt test takers to invest effort to avoid failure, and may facilitate rigid learning strategies (like simple rehearsal), thus exerting task-dependent effects on motivation and achievement which may be quite ambivalent (Eysenck and Calvo, 1992; Meinhardt and Pekrun, 2003; Pekrun *et al.*, 2002a, b; Zeidner, 1998).

Finally, the theory also assumes that achievement may exert reverse effects on achievement-related emotions. Feedback on achievement can affect perceived control as well as subjective values of achievement, which in turn underlie the development of achievement emotions, according to the theory. Addressing effects of emotions on learning/achievement as well as reverse effects of performance feedback on emotions, control-value theory assumes that emotions and achievement are linked by feedback loops implying reciprocal causation.

EXPLORATORY ANALYSIS OF EMOTIONS EXPERIENCED IN EXAM CONTEXTS (STUDIES 1 AND 2)

To ensure that prevalent test emotions were included in the TEQ, exploratory analysis was used to gather base rate information on the relative frequencies of various test emotions. Furthermore, we wanted to obtain empirical information on the component structures of test emotions, and on the phase of the exam process in which they are experienced (before, during, or after the exam). Finally, we also used the self-report statements delivered by study participants to construct item pools for the TEQ. Study 1 was part of an investigation of bio-behavioral relations between students' feelings, psychophysiological processes, and behaviors in exams (Spangler *et al.*, 2002). In this study, we used an open-ended interview format to analyze the emotions experienced by university students during an important final exam. In Study 2, in contrast, we used a semi-structured questionnaire to explore the emotions experienced by students at different phases of the exam process; that is, before, during, and after an exam.

Method

Study 1. Immediately after an important oral exam, we interviewed $N=26$ teacher training students at the University of Regensburg about the occurrence, structures, and correlates of the emotions they had experienced during the exam (15 females and 11 males, mean age 25.2 years, range 22–31 years). The 30-minute oral exam was one of the students' final university exams. The grades awarded counted towards their final score, which largely determined their chances of getting a job as a teacher, meaning that it was a high-stakes exam. The exam was videotaped, and the interview was based on video reconstruction of students' experiences. Using event sampling of episodes within the exam, participants were asked to watch the video and to stop it whenever it showed an episode during which they had experienced a significant emotion or thought. Using an open-ended interview technique, the participants were asked to report these emotions or thoughts, as well as their bodily reactions, coping behavior, and interpretation of the situation. The number of self-selected episodes per subject

ranged from 5 to 16 (with a mean of 7.7 episodes); the total number of episodes was 185. The interviews were tape recorded, transcribed, and analyzed using a classification system for exam emotions and their correlates adapted from Pekrun (1992b). The system produced 51 categories of emotions, which were then combined into the summary categories displayed in Table I (e.g., categories for nervousness, anxiety, and panic were combined into one category of "anxiety").

Study 2. An exploratory, semi-structured questionnaire was used to assess frequencies, components, and subjective correlates of emotions experienced before, during, and after exams. The sample consisted of $N = 186$ students enrolled in different courses in the humanities, psychology, or law at the University of Regensburg (122 females, 64 males, mean age 24.0 years, $SD = 3.2$). The questionnaire was administered immediately after the students took an important oral ($n = 98$) or written ($n = 88$) exam and asked them to describe the feelings, thoughts, and behaviors they experienced before the exam, during the exam, and afterwards. The questionnaire consisted of three sections, one for each time interval. Each of the three sections contained two parts. In the first part, a list of emotions was presented (see Table I), and participants were asked to indicate which of these emotions they had experienced during the respective time interval. In the second part, they were asked to give a qualitative description of the contents and correlates of one of these emotions. The order of emotions on the list and the emotion to be described were randomized across participants.

TABLE I Percentages of Different Test Emotions Experienced in Exam Contexts

Emotion	Study 1					Study 2			
	Start	Middle	End	Total 1	Total 2	Before	During	After	Total
<i>Positive</i>									
Joy	0.6	3.5	2.8	2.5	38	11.6	33.1	63.6	35.9
Hope	4.5	4.4	4.2	4.3	50	54.2	64.2	64.2	60.8
Pride	1.3	1.3	0.7	1.1	15	25.2	50.3	64.2	46.4
Relief	8.3	10.0	11.2	9.8	77	21.9	64.2	88.7	58.0
Satisfaction	—	0.9	1.4	0.8	15	—	—	—	—
Gratitude	—	—	—	—	—	5.2	25.2	45.7	25.2
Admiration	—	—	—	—	—	13.5	12.6	22.5	16.2
Feeling of security	4.5	6.1	3.5	4.9	54	35.5	58.3	57.6	50.3
Other positive	10.2	10.4	11.2	10.6	73	—	—	—	—
<i>Negative</i>									
Anger	9.6	11.7	16.8	12.5	77	14.8	18.5	15.2	16.2
Anxiety	34.4	20.4	16.8	23.6	96	80.6	59.6	21.2	54.0
Shame	1.9	3.0	2.8	2.6	31	7.1	14.6	11.3	10.9
Hopelessness	2.5	4.4	2.1	3.2	42	25.2	21.2	11.3	19.3
Sadness	—	—	—	—	—	26.5	11.9	18.5	19.0
Disappointment	0.6	2.2	1.4	1.5	19	3.2	25.8	25.8	18.2
Envy	—	—	—	—	—	6.5	4.6	5.3	5.5
Other negative	12.7	13.9	16.1	14.1	69	—	—	—	—
<i>Neutral</i>									
Surprise	8.9	6.1	6.3	7.1	42	1.9	39.1	23.2	21.2
Indifference	0.0	1.7	2.8	1.5	23	—	—	—	—

Note. Study 1: Start, middle, end of exam: Percentage of total number of emotions reported per time interval. Total 1: Percentage of total number of emotions reported for all time intervals. Total 2: Percentage of participants reporting the emotion. Study 2: Percentage of participants reporting the emotion.

Results and Discussion

Table I displays the relative frequencies of the different categories of emotions, broken down by the respective phase of the exam (start, middle, and end of the exam in Study 1; before, during, and after the exam in Study 2).

In Study 1, 530 emotions were reported for the 185 exam episodes selected by the participants. Of these emotions, 180 were positive, 305 negative, 38 related to surprise, and seven to the neutral category of "feelings of indifference." Anxiety accounted for 23.6% of students' reported emotions, implying that the remaining 76.4% was made up by other emotions. Concerning negative emotions, anxiety was reported most often, but anger was also mentioned quite frequently, as were shame/guilt and hopelessness. As argued by Zeidner (1998, p. 294), it can be assumed that *anger* may be aroused when an exam situation is perceived as frustrating important personal goals. The anger reported by our participants related to the examiner, the questions asked, and the examiner's reactions to students' answers (*other-directed anger*), or to the student her- or himself (*self-directed anger*, e.g., anger about one's own lack of preparation). In the last phase of the exam, anger was reported as often as anxiety. Interestingly, the primary emotion of *sadness* was not reported at all. This finding is congruent with evidence reported by Scherer *et al.* (1986) indicating that frequencies of (pure) sadness in achievement situations are low. Apparently, achievement and exam situations are more likely to induce specific variants of this emotion (e.g., hopelessness) than sadness *per se*. As to positive emotions, hope, relief, and joy were reported frequently in Study 1. Students' feelings of *joy* during the exam mainly related to positive competence experiences (being able to answer questions), *hope* about getting questions they could answer, and *relief* at having been able to give an appropriate answer.

In Study 2, the average number of emotions reported per time interval was 4.9 ($M = 3.6, 5.3, \text{ and } 5.6$ for the intervals before, during, and after the exam, respectively). As can be seen from Table I, positive emotions were not reported any less often than negative emotions. However, the relative frequencies differed according to the time interval. *Before* the exam, negative emotions were more frequent than positive emotions: 80.6 and 54.2% of the sample reported having experienced the prospective, uncertainty-based emotions of anxiety and hope, respectively, whereas other emotions were reported less frequently. Nevertheless, a number of additional emotions did seem to play a role before the exam, also including hopelessness. *During* the exam, anxiety and hope continued to play a major role. However, the outcome-contingent emotions of relief, pride, disappointment, and shame now also were reported frequently, largely based on perceived success or failure in responding to the exam questions. *After* the exam, positive emotions outweighed negative ones. Relief replaced anxiety as the emotion mentioned most frequently, and joy, pride, continuing hope, and gratitude were also reported often. The negative emotions mentioned included sadness, disappointment, anger, and shame, as well as anxiety. Participants' reports indicate that the hope and anxiety they experienced after the exam related to the imminent feedback about their results.

The average number of emotions reported per time interval and the absolute frequencies of different emotions were higher in Study 2 than in Study 1. This can probably be attributed to the closed response format of that part of the questionnaire which was used to obtain frequency estimates in Study 2, in contrast to the open-ended

format of the interviews in Study 1, which relied on students' active recollections of their emotions. With respect to the relative frequencies of different emotions and the multiple nature of students' emotional experiences, findings from Studies 1 and 2 converged. An exception were the results on anger (when the relative frequencies of the different emotions are arranged in rank order, anger took second position in Study 1, but only position 12 for the "during the exam" interval in Study 2). This may have been due to differences in the two student samples, or to differences in the exams sampled in the two studies (oral face-to-face exams facilitating social anger in Study 1, both oral and written exams in Study 2).

Conclusions

In sum, results of Studies 1 and 2 show that students experience a variety of positive and negative emotions during exams. Anxiety was the single emotion reported most often in both studies, yet reports of anxiety accounted for only part of the emotions experienced. It seems that, even in the stressful, critical life event of an important examination, anxiety is by no means the only frequent emotion. Positive emotions like hope, relief, and joy can play a major role as well. Concerning negative emotions, anger, shame, and hopelessness may be important, in addition to or instead of anxiety.

CONSTRUCTION OF THE TEQ SCALES (STUDY 3)

The findings of Studies 1 and 2 suggest that any more complete measure of test emotions should contain more than just a scale measuring test anxiety. In Study 3, we made an initial attempt to operationalize a number of major test emotions by constructing self-report scales to measure test-related joy, hope, relief, anger, anxiety, and hopelessness (Table II), and analyzed the item and scale characteristics of this first version of the TEQ.

Selection of emotions. The emotions included in the TEQ were selected according to the following criteria. (a) *Frequency:* The TEQ aims at representing test-related emotions which are experienced often and constitute typical reactions to tests and exams. Based on the frequency findings of our exploratory studies, joy, hope, relief, anger, and anxiety were included. (b) *Inclusion of emotions only:* The TEQ measures test-related emotions. It does not attempt to assess test-related subjective states more generally (as can be done by mood scales). Therefore, non-emotional states like "feelings of security" reported in our exploratory studies were not included. (c) *Relevance for important outcomes:* The questionnaire seeks to capture those test emotions which can be regarded as relevant for achievement, health, and personality development. Test hopelessness was thus added because of its theoretical importance for achievement and health, irrespective of low frequency scores in our exploratory data. Furthermore, pride and shame were also included in the revised TEQ (see below, Studies 4 and 5) because of their impact on motivation and achievement, and their central status in achievement motivation theory (Heckhausen, 1991). (d) *Taxonomic representativeness:* The resulting list of emotions ensures that the TEQ represents all four quadrants of the valence \times activation classification discussed above (positive

TABLE II Item Statistics and Reliabilities of the TEQ Scales

	Study	Number of items	Average item statistics (median)				α
			<i>M</i>	<i>SD</i>	r_{it}	r_{dt}^a	
Joy	3	32	1.82	1.02	0.41	0.30	0.89
	4	19	2.08	1.02	0.52	0.27	0.91
	5	23	1.96	0.98	0.51	0.27	0.90
	6	10	1.83	0.58	0.45	0.25	0.78
Hope	3	16	2.53	1.05	0.31	0.24	0.70
	4	18	2.44	0.91	0.57	0.31	0.90
	5	16	2.38	0.92	0.55	0.29	0.89
	6	8	2.24	0.62	0.52	0.31	0.80
Pride	4	18	1.61	0.99	0.62	0.27	0.93
	5	16	1.77	0.98	0.61	0.25	0.92
	6	10	2.13	0.65	0.58	0.29	0.86
Relief	3	29	2.51	1.14	0.41	0.21	0.87
	4	26	1.75	0.97	0.50	0.18	0.88
	5	14	2.84	0.98	0.58	0.15	0.89
	6	7	2.52	0.71	0.47	0.11	0.75
Anger	3	28	0.98	0.86	0.50	0.22	0.91
	4	19	1.18	0.94	0.58	0.20	0.91
	5	17	1.14	0.93	0.55	0.20	0.89
	6	10	1.34	0.73	0.57	0.28	0.86
Anxiety	3	31	1.53	1.13	0.55	0.25	0.93
	4	31	1.75	1.08	0.52	0.28	0.92
	5	31	1.61	1.07	0.58	0.30	0.94
	6	12	2.02	0.83	0.61	0.34	0.90
Shame	4	20	0.97	0.96	0.53	0.23	0.90
	5	19	0.97	0.99	0.60	0.26	0.93
	6	10	1.19	0.75	0.60	0.33	0.87
Hopelessness	3	32	0.97	0.92	0.53	0.35	0.94
	4	24	0.67	0.77	0.61	0.30	0.94
	5	21	0.59	0.78	0.64	0.31	0.94
	6	11	1.01	0.77	0.69	0.38	0.92

Note. Pride and shame scales: Studies 4–6 only. Range of item values: 0–4.

^aMedian of correlations (absolute values) with other emotion scales.

activating emotions: joy, hope, and pride; positive deactivating: relief; negative activating: anger, anxiety, shame; negative deactivating: hopelessness).

Item construction: Representing the contents of emotion constructs. Based on our exploratory findings and the self-report statements produced by the participants in Studies 1 and 2, an initial item pool of more than 500 items was constructed. Items for the first version of the TEQ were selected from this pool. For the affective, cognitive, and physiological components of test anxiety, we used the tension, worry, and bodily reactions subscales of Sarason's (1984) Reactions to Tests questionnaire (RTT). New items for the motivational component of test anxiety (relating to tendencies to escape test situations) were added.

In constructing and selecting items, we aimed at establishing content validity by covering all domains of the test emotions measured by the TEQ. Specifically, we sought to represent the different emotion components (affective, cognitive, physiological, motivational) and elements of these components as defined by our taxonomy of test

emotion components, as well as the time reference of different test emotions (before, during, and after test-taking). Furthermore, we wanted to construct items such that they can be used to measure both trait and state versions of test emotion constructs. Therefore, we avoided the use of frequency terms ("often," "rarely," etc.). Finally, in order to obtain valid indicators of respondents' recollections of emotions, we tried to use everyday language by drawing on the self-report statements produced by the participants in our exploratory studies. Items were worded as statements followed by five-point Likert agreement ratings ("strongly disagree" to "strongly agree").

Situation-reaction format of scales. Many test anxiety scales are plagued by redundancy, with time references being repeated in each item ("Before an exam, I feel . . .," "Before an exam, my stomach . . ."). In order to reduce redundancy, and to help respondents to retrieve situation-related information from their autobiographical memories more easily, we used a situation-reaction format to structure the TEQ. Items do not contain time reference terms. Instead, they are grouped into three situational blocks pertaining to the intervals before, during, and after test-taking. Each block is preceded by a short, general description defining this situational reference (e.g., "The following questions pertain to feelings you may experience *before* taking a test or an exam. Please indicate how you feel, typically, before taking a test or an exam"). Within blocks, items measuring different emotions and components within emotions are presented in a balanced order, as are positively versus negatively worded items, in order to reduce the effects of response sets.

In a sample of university students, we analyzed item and scale characteristics of the first version of the TEQ, as well as their relations to demographic variables, personality, learning, achievement, and perceived health.

Method

The TEQ scales were administered to $N=150$ students enrolled in different courses at the University of Regensburg (90 females, 60 males, mean age 24, range 19–44 years). Participation in the study was voluntary, and the questionnaire and achievement data were gathered anonymously. In addition to the TEQ, we included self-report scales on general trait anger (10 items) and anxiety (nine items) from Spielberger's State-Trait Personality Inventory (STPI; German version, Schwarzer, 1986), general self-esteem (12 items; Schwarzer, 1986), study-related self-efficacy (seven items; Jerusalem and Schwarzer, 1986), study interest (16 items; Winteler *et al.*, 1991), effort when studying (seven items; Winteler *et al.*, 1991), irrelevant thinking when studying (six items; Winteler *et al.*, 1991), irrelevant thinking when taking tests (nine items; German version of the Irrelevant Thinking subscale of the Reactions to Tests questionnaire, Quast *et al.*, 1986), and perceived health problems (subscales with two, three, and two items on cardiovascular, stomach, and back problems, respectively; Zerssen, 1976). Students' average grades in their mid-studies exams were included in the analysis as an indicator of achievement. Students at German universities usually take mid-studies exams at the end of their second year of studies. Admission to the next phase of the academic program is contingent on passing these exams.

Results and Discussion

Score distributions, item statistics, and reliability. The scales showed good psychometric properties, including sufficient variation of scores at both item and scale levels, and internal consistencies above $\alpha = 0.87$ (Table II). The reliability coefficient for the hope scale was an exception ($\alpha = 0.70$), probably due to the complex wording of some of the items.

Internal construct validity. Correlational analysis showed that 25 of the 36 correlations between the subscales of each scale were in the range of $0.35 < r < 0.65$, thus implying discriminability of the emotion components, as well as sufficient overlap justifying their combination into emotion scales. Concerning the interrelations of the emotion scales, most of the correlations were low to medium, indicating divergent validity of the scales (see the entries for Study 3 in Table III). As expected, both the scales for positive emotions and the scales for negative emotions showed positive intercorrelations. Correlations between joy and hope, on the one hand, and negative emotions, on the other, were low, corroborating that positive versus negative test emotions do constitute more than just one bipolar dimension of positive versus negative feelings towards exams. Finally, correlations between test relief and negative test emotions were moderately positive. Relief can be regarded as a tension-reducing emotion linked to the termination of subjectively negative states, implying that it should be related positively to the experience of preceding negative emotions.

Divergent validity of the test anxiety and test hopelessness scales. An exception to the rule of divergent validity was the high correlation between test anxiety and test hopelessness. Since the affective, cognitive, and physiological subscales of the test anxiety scale were derived from Sarason's RTT, this finding also implies that the RTT might lack divergent validity. More generally, since the RTT items are fairly representative of contemporary test anxiety scales, high correlations with measures of hopelessness would imply that these scales might still be in danger of "measuring more than they denote" (Nicholls, 1976).

To gain a deeper understanding of the problems entailed, we correlated the four subscales of the test anxiety scale as well as their individual items with the test hopelessness subscales. The correlations between the affective, cognitive, physiological, and motivational anxiety and hopelessness subscales were $r = 0.63, 0.80, 0.59,$ and $0.76,$ respectively, implying a satisfactory level of differentiation in the affective and physiological components of the two emotions (feeling anxious versus hopeless; physiological activation versus deactivation), but a lack of discrimination as regards the cognitive and motivational components. Concerning *cognitive* components of test-related anxiety and hopelessness, both emotions may be characterized by worries about impending failure and its consequences, the difference being the relative amount of subjective certainty (uncertainty of negative events in anxiety, certainty in hopelessness). It would seem that many traditional worry items in test anxiety questionnaires fail to take this difference into account. For example, items like the RTT worry item "Before taking a test, I worry about failure" do not specify the amount of subjective certainty and may therefore be regarded as measuring both anxiety-related and hopelessness-related worry. Consequently, in our analysis this item was as highly correlated with the test anxiety scale as it was with the test hopelessness scale ($r = 0.30/0.32$). Similar difficulties may arise with respect to the *motivational* component that

TABLE III Intercorrelations of the TEQ Scales

	Study	Emotion						
		Joy	Hope	Pride	Relief	Anger	Anxiety	Shame
Hope	3	0.61***	—					
	4	0.77***	—					
	5	0.68***	—					
	6	0.70***	—					
Pride	4	0.85***	0.72***	—				
	5	0.83***	0.62***	—				
	6	0.71***	0.68***	—				
Relief	3	0.47***	0.55***	—	—			
	4	0.44***	0.24***	0.31***	—			
	5	0.28***	0.03	0.26***	—			
	6	0.10*	-0.01	0.25***	—			
Anger	3	0.14	0.21**	—	0.46***	—		
	4	-0.12*	-0.26***	-0.05	0.13*	—		
	5	-0.17**	-0.28***	-0.08	0.15*	—		
	6	-0.25***	-0.36***	-0.20***	0.08	—		
Anxiety	3	-0.03	0.05	—	0.40***	0.57***	—	
	4	-0.29***	-0.45***	-0.32***	0.31***	0.55***	—	
	5	-0.33***	-0.48***	-0.23***	0.42***	0.46***	—	
	6	-0.39***	-0.48***	-0.29***	0.35***	0.57***	—	
Shame	4	-0.15*	-0.33***	-0.13*	0.12*	0.51***	0.68***	—
	5	-0.16*	-0.36***	-0.07	0.25***	0.47***	0.67***	—
	6	-0.30***	-0.43***	-0.37***	0.06	0.64***	0.66***	—
Hopelessness	3	0.00	-0.07	—	0.28***	0.53***	0.83***	—
	4	-0.35***	-0.58***	-0.35***	0.00	0.55***	0.68***	0.62***
	5	-0.35***	-0.61***	-0.27***	0.16*	0.53***	0.70***	0.59***
	6	-0.39***	-0.53***	-0.40***	-0.02	0.73***	0.67***	0.79***

Note. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

tends not to be included in contemporary test anxiety scales. Conceptually, anxiety may be characterized by motivation to avoid failure, hopelessness by low motivation and resignation. Operationally, however, it seemed to be difficult to reflect these differences in the self-report items constructed for the first version of the TEQ.

Relations to demographic variables. Due to the restricted age range of the present sample, the relations of the test emotions scores to age were near zero and not significant. In line with previous research (Hembree, 1988), scores for test anxiety were higher among female than male students, although the difference was small and failed to reach significance (see Table IV). Congruent with the gender difference in anxiety scores, mean scores for relief and hopelessness also were higher for female students. For joy, hope, and anger, no relevant differences emerged.

Relations to personality. As situation-specific trait emotions, test emotions imply habitual individual tendencies to react to test situations by experiencing specific emotions. Scales measuring such emotions should thus be linked to measures of general trait emotions (i.e., tendencies to experience specific emotions in a variety of situations). Correlations of test anger, test anxiety, and test hopelessness were $r = 0.42$, 0.38 , and 0.31 , respectively, with *general trait anger*, and $r = 0.21$, 0.65 , and 0.60 , respec-

tively, with *general trait anxiety* ($p < 0.05$ for all coefficients), thus implying that the test anger and anxiety scales correlated significantly with general trait anger and anxiety, respectively. Furthermore, test anxiety and hopelessness correlated negatively, and test joy positively, with students' *general self-esteem* (see the entries for Study 3 in Table V). For test anxiety, this finding is in line with the results of previous studies (Hembree, 1988). However, most correlations with general self-esteem were weaker than the correlations with domain-specific, study-related academic self-efficacy (Table V).

Relations to academic learning and achievement. According to Pekrun's control-value theory summarized above, control-related appraisals should be major determinants of achievement-related emotions, and motivation, learning strategies, and cognitive resources should be major variables mediating the impact of these emotions on achievement. In line with these assumptions, students' study-related self-efficacy correlated positively with test joy and hope, and negatively with test anger, anxiety, and hopelessness (Table V). Concerning variables mediating achievement effects, test

TABLE IV Test Emotion Scores for Male and Female Students

	Study	Males		Females		<i>t</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Joy	3	59.35	16.22	57.57	14.94	-0.67
	4	61.84	15.72	56.93	14.61	-2.36*
	5	46.71	13.88	44.41	11.93	-1.23
	6	18.91	6.20	17.97	5.89	-1.50
Hope	3	40.95	6.22	40.20	6.34	-0.69
	4	46.81	10.25	42.75	9.87	-2.99**
	5	40.29	9.85	36.95	8.49	-2.54*
	6	18.59	5.15	17.50	4.74	-2.11
Pride	4	31.27	11.64	28.08	11.91	-1.97
	5	30.07	10.96	27.45	10.02	-1.73
	6	21.85	6.61	20.99	6.43	-1.27
Relief	3	69.89	14.30	74.86	14.35	2.01*
	4	42.71	11.79	46.73	9.60	2.83**
	5	37.12	9.13	41.03	8.59	3.06**
	6	17.27	5.03	17.86	4.42	1.15
Anger	3	27.03	15.63	27.57	13.78	0.22
	4	19.91	10.33	23.76	11.27	2.57*
	5	18.00	9.86	20.14	9.68	1.50
	6	13.64	7.78	13.07	6.94	-0.74
Anxiety	3	43.57	16.65	49.66	21.22	1.81
	4	48.08	17.62	57.33	18.93	3.39**
	5	49.51	17.01	62.28	19.92	4.55**
	6	22.08	9.53	25.60	10.09	3.39**
Shame	4	15.81	9.92	21.09	11.66	3.46**
	5	12.72	9.96	21.05	12.37	4.83**
	6	11.11	7.53	12.34	7.50	1.56
Hopelessness	3	28.69	15.43	32.43	18.31	1.26
	4	11.81	10.16	17.91	12.39	3.53**
	5	9.88	8.70	14.12	11.36	3.30**
	6	10.61	8.19	11.29	8.56	0.77

Note. *n* (males/females): Study 3: *n* = 60/90; Study 4: *n* = 84/168; Study 5: *n* = 71/151, Study 6: *n* = 155/234.

* $p < 0.05$. ** $p < 0.01$.

joy and hope correlated positively with study interest and effort, whereas the correlations for hopelessness were significantly negative. This is in line with our theoretical assumptions on the effects of positive activating and negative deactivating emotions. Test anxiety correlated positively with task-irrelevant thinking, but so did hopelessness and, to a lesser extent, anger and shame. In test anxiety research, it has often been assumed that there is a specific link between test anxiety and irrelevant thinking, or that irrelevant thinking should be regarded as part of the test anxiety construct (Sarason, 1984; Zeidner, 1998). The present data imply that there may be no unique relationship, but that other negative test emotions may produce equally intruding, task-irrelevant thoughts.

Finally, five of the six scales correlated significantly with students' mid-studies achievement. The correlations for test joy and hopelessness were higher than the correlation for test anxiety, in line with our assumptions that positive activating and negative deactivating emotions have consistent effects on achievement, in contrast to the ambivalence of test anxiety effects. Findings thus suggest that test emotions other than anxiety may be just as important for students' achievement as test anxiety has been shown to be (Zeidner, 1998). However, caution should be exerted in interpreting these findings because of their exploratory, cross-sectional nature, and the preliminary status of the scales used in this study.

Relations to health complaints. Test anger, anxiety, and hopelessness correlated positively and significantly with perceived back, cardiovascular, and stomach problems (Table VI). Correlations for relief were positive as well, probably due to the close relation between negative emotions and subsequent relief, and less to direct causal links between relief and health problems. Interestingly, the correlations for test emotions were no less strong than the correlations for general trait anger and anxiety. These findings suggest that students' emotional experiences relating to taking tests and exams significantly affect not only their academic performance, but also their health. Again, correlations should be interpreted with caution because of their cross-sectional nature.

Conclusions

Findings imply that our first attempt at constructing psychometric scales to measure a number of positive and negative test emotions was successful. Their psychometric quality, internal construct validity, and external relations suggest that the six scales are reliable and show structural as well as external validity. However, the reliability of the test-related hope scale was not yet optimal. A second problem was the lack of divergent validity of the test anxiety and test hopelessness scales. These two weaknesses were addressed when revising the scales (Studies 4 and 5).

REVISION OF THE TEQ SCALES (STUDIES 4 AND 5)

In Studies 4 and 5, revised and final versions of the TEQ scales were developed and validated (see Appendix for sample items from the final TEQ scales). In Study 4, the TEQ was revised based on the findings of Study 3. In addition, scales measuring test-related pride and shame were developed and included. In Study 5, the TEQ scales were revised once more to produce final versions. Parallel procedures were used for item and

TABLE V Correlations of Test Emotions with Self esteem, Variables of Learning, and Academic Achievement

	Study	General self-esteem	Academic self-efficacy	Academic control	Interest	Effort	Elaboration	Rehearsal	Irrelevant thinking when learning	Irrelevant thinking when taking tests	Self-regulation of learning	Achievement
Joy	3	0.22**	0.26**	—	0.30***	0.10	—	—	0.03	-0.02	—	0.33***
	4	—	0.44***	0.27***	0.27***	0.41***	0.32***	0.07	-0.27***	-0.11	0.41***	0.24*
	5	—	0.49***	0.31***	0.40***	0.31***	0.44***	-0.02	-0.32***	-0.26***	0.50***	0.36**
	6	0.36***	0.50***	0.20***	0.35***	0.37***	0.38***	0.23***	—	—	0.38***	0.26***
Hope	3	0.13	0.21**	—	0.12	0.23**	—	—	0.11	-0.16*	—	0.17*
	4	—	0.55***	0.43***	0.30***	0.44***	0.29***	-0.01	-0.30***	-0.22***	0.42***	0.35***
	5	—	0.60***	0.39***	0.31***	0.40***	0.33***	-0.03	-0.37***	-0.28	0.42***	0.37**
	6	0.55***	0.60***	0.35***	0.37***	0.44***	0.40***	0.18***	—	—	0.51***	0.25***
Pride	4	—	0.54***	0.33***	0.21***	0.32***	0.23***	-0.01	-0.22***	-0.12*	0.36***	0.23*
	5	—	0.49***	0.25***	0.26***	0.30***	0.35***	0.08	-0.26***	-0.24***	0.41***	0.26*
	6	0.51***	0.56***	0.33***	0.32***	0.42***	0.44***	0.31***	—	—	0.49***	0.34***
Relief	3	-0.05	-0.10	—	0.00	-0.02	—	—	0.24**	0.08	—	0.37***
	4	—	-0.04	0.03	0.08	0.26***	0.20**	0.16**	0.02	-0.16**	0.19**	0.01
	5	—	-0.13**	-0.10*	0.02	0.12*	0.06	0.22***	0.13**	-0.17***	0.14**	-0.03
	6	0.04	-0.10*	0.17***	0.16**	0.06	0.22***	0.20***	—	—	0.18***	0.15*
Anger	3	-0.12	-0.24**	—	-0.09	-0.25**	—	—	0.27***	0.24**	—	-0.14
	4	—	-0.40***	-0.45***	-0.22***	-0.14*	-0.08	0.21***	0.36***	0.26***	-0.15*	-0.37***
	5	—	-0.30***	-0.33***	-0.25***	-0.10	-0.02	0.22**	0.39***	0.37***	-0.13*	-0.33**
	6	-0.53***	-0.40***	-0.54***	-0.32***	-0.37***	-0.29***	-0.03	—	—	-0.30***	-0.32***
Anxiety	3	-0.38***	-0.53***	—	-0.06	-0.34***	—	—	0.43***	0.43***	—	-0.15
	4	—	-0.51***	-0.40***	-0.11	-0.10	0.01	0.14*	0.32***	0.15*	-0.17**	-0.42***
	5	—	-0.60***	-0.43***	-0.16**	-0.06	-0.19***	0.29***	0.38***	0.15**	-0.30***	-0.24*
	6	-0.46***	-0.38***	-0.30***	-0.13	-0.28***	-0.11*	0.12*	—	—	-0.28***	-0.14*
Shame	4	—	-0.39***	-0.38***	-0.13*	-0.14*	0.01	0.15*	0.33***	0.21***	-0.18**	-0.44***
	5	—	-0.42***	-0.31***	-0.10*	-0.15**	-0.11*	0.18***	0.34***	0.14**	-0.30***	-0.24*
	6	-0.58***	-0.43***	-0.47***	-0.19***	-0.38***	-0.25***	-0.01	—	—	-0.37***	-0.37***
Hopelessness	3	-0.35	-0.52***	—	-0.20*	-0.41***	—	—	0.45***	0.43***	—	-0.30***
	4	—	-0.57***	-0.53***	-0.28***	-0.25***	-0.12*	0.14*	0.36***	0.30***	-0.31***	-0.49***
	5	—	-0.54***	-0.38***	-0.26***	-0.23***	-0.21***	0.10*	0.36***	0.36***	-0.30***	-0.31**
	6	-0.65***	-0.51***	-0.60***	-0.33***	-0.45***	-0.35***	-0.04	—	—	-0.41***	-0.34***

Note. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

TABLE VI Correlations of Test Emotions and General Trait Emotions with Perceived Health Problems

	Study	Type of problem			
		Cardiovascular	Back	Stomach	Sleep
<i>Test emotions</i>					
Joy	3	-0.10	-0.04	0.13	-
	5	-0.07	0.00	-0.07	-0.10
Hope	3	0.04	0.06	0.06	-
	5	-0.23**	-0.13	-0.23**	-0.19**
Pride	5	-0.10	-0.02	-0.10	-0.08
Relief	3	0.21*	0.22*	0.24*	-
	5	0.14*	0.18**	0.14*	0.29**
Anger	3	0.25*	0.27**	0.37**	-
	5	0.25**	0.08	0.28**	0.26**
Anxiety	3	0.25	0.28**	0.37**	-
	5	0.30**	0.16*	0.37**	0.42**
Shame	5	0.33**	0.18**	0.33**	0.35**
Hopelessness	3	0.31**	0.20**	0.39**	-
	5	0.29**	0.10	0.34**	0.36**
<i>General trait emotions</i>					
Anger	3	0.28**	0.32**	0.36**	-
Anxiety	3	0.18	0.15	0.32**	-

Note. * $p < 0.05$. ** $p < 0.01$.

scale analysis in Studies 4 and 5. As in Study 3, these procedures pertained to distributions of item and scale scores, scale reliabilities, convergent versus divergent item validities, and intercorrelations of scales, as well as their external relations to demographic variables, academic learning and achievement, and perceived health problems.

Method

The samples consisted of $N = 252$ students in Study 4 (168 females, 84 males, mean age = 24.1, $SD = 3.9$ years) and $N = 222$ students in Study 5 (151 females, 71 males, mean age 23.1 years, $SD = 3.3$ years). Students were enrolled in different courses at the University of Regensburg. Participation in the studies was voluntary, and the questionnaire was answered anonymously. In both studies, the first part of the questionnaire contained the TEQ (revised version in Study 4, final version in Study 5). The second part included questions on demographic variables as well as the following self-report scales: (a) *Study-related control beliefs*: academic self-efficacy (Jerusalem and Schwarzer, 1986; 7 items; $\alpha = 0.83/0.80$ in Studies 4 and 5); perceived academic control (Perry, 1991; 10 items; $\alpha = 0.74/0.79$); (b) *motivation*: study interest (Schiefele *et al.*, 1993; nine items, $\alpha = 0.83/0.84$); effort (Wild and Schiefele, 1994; eight items, $\alpha = 0.75/0.78$); (c) *learning strategies*: elaboration (Wild and Schiefele, 1994; 10 items, $\alpha = 0.83/0.80$); rehearsal (Wild and Schiefele, 1994; seven items; $\alpha = 0.74/0.74$); these two scales are based on Pintrich's Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich *et al.*, 1991); (d) *cognitive resources*: irrelevant thinking when learning

(Wild and Schiefele, 1994; six items, $\alpha = 0.92/0.92$); irrelevant thinking when taking tests (Quast *et al.*, 1986; seven items, $\alpha = 0.95/0.90$); (e) *self-regulation of learning* (scale on self-regulation of the goals, strategies, and control of learning; Titz, 2001; seven items, $\alpha = 0.72/0.69$); (f) *health complaints*: cardiovascular, stomach, back, and sleep problems (Zerssen, 1976; two/three/two/two items). Finally, as in Study 3, students' average grades in their mid-studies exams were used to assess their *academic achievement*.

Results and Discussion

Score distributions, item statistics, and reliability. The distributions of the item and scale scores of the revised and final scales indicated that there continued to be sufficient variation. Most of the items and scales also showed sufficient symmetry. An exception was the positive skewness (1.37) of the hopelessness scale. The asymmetry of scores for this scale should not be interpreted as a methodological flaw, however. For as extreme an emotion as hopelessness, low base rates are to be expected and were indeed found in Studies 1 and 2 (see above). Constructing hopelessness items in such a way that distributional symmetry is achieved in normal student populations might distort reality. For both the revised and the final scales, item-total correlations were high, as were the reliabilities ($0.88 < \alpha < 0.95$ in both studies; Table II). In contrast to the first version of the TEQ, the revised and final test hope scales now also showed good internal reliability.

Convergent versus divergent item validities. One aim of scale revision was to modify items so that neighboring emotions like test-related anxiety, hopelessness, and shame can be differentiated. Table II displays summary statistics for divergent item validities (comparison of part-whole corrected item-total correlations versus correlations with the other emotion scales). As may be seen from Table II, compared to the first version of the TEQ, divergent validities were improved for the joy, hope, relief, anger, and hopelessness scales by revising the scales. For the final TEQ scales, divergent item validities proved to be sufficiently high for most pairs of scales.

There were exceptions, however, which persisted after selecting items according to their divergent validity. Most importantly, even for the final scales, the mean correlations of anxiety items with the hopelessness and shame scales were $r = 0.43$ and 0.44 (the mean item-total correlation being $r = 0.58$), and the mean correlations of joy items with the hope and pride scales were $r = 0.40$ and 0.47 (mean item-total correlation of $r = 0.51$). Apparently, as individual differences constructs, test-related anxiety, shame, and hopelessness are so closely related that trait items measuring these constructs show substantial correlative overlap: Individuals prone to experience test anxiety also tend to experience shame and hopelessness in test and exam situations. In contrast, the correlations between state test anxiety, shame, and hopelessness proved to be lower (Pekrun *et al.*, in prep.; e.g., in a sample of university students, the correlation between state anxiety and state hopelessness during an exam was $r = 0.55$). Similarly, the individual differences constructs of trait test-related joy, hope, and pride also seem to be closely linked. Hope and pride can be regarded as specific variants of the primary emotion of joy, implying overlap between these constructs.

Interrelations of scales. Similar to the first version of the TEQ scales, the scales for positive test emotions correlated positively, as did the scales for negative test emotions

(Table III). The correlations between positive and negative emotions were negative, with the exception of relief, which again correlated positively with both positive and negative emotions. The correlations of relief with joy and pride are probably due to shared valence and partially shared situational antecedents, whereas the relations with anxiety may be caused by the relaxing properties of relief presupposing preceding affective tension. Overall, the pattern of correlations implied sufficient divergent validity of the scales. In contrast to the correlations of the first versions of the TEQ scales, the intercorrelations for the anxiety, shame, and hopelessness scales were also well below the respective reliabilities. An exception from divergent validity were the high correlations between test-related joy and pride. As argued above, these correlations are congruent with the overlapping nature of the two constructs of joy and pride.

Relations to demographic variables. The correlations of emotion scores with students' age were again near zero, and none of them reached the 1% level of significance. However, there were significant differences between the genders. In Studies 4 and 5, female students had significantly higher scores for test-related anxiety, shame, and hopelessness, as well as lower scores for relief (see Table IV).

Relations to academic learning and achievement. The findings regarding relations with learning and achievement confirmed the results of Study 3. Test-related joy, hope, and pride were positively related to self-efficacy and perceived academic control, whereas the correlations for negative test emotions were negative. In addition, there were significant, consistent relations between positive activating emotions (joy, hope, and pride) as well as deactivating hopelessness, on the one hand, and study-related interest, effort, elaboration of learning material, irrelevant thinking, and self-regulation of learning, on the other. To the extent that these relations represent the effects of emotions on performance (rather than reverse effects of performance on emotions), they imply that positive activating test emotions may be beneficial for students' motivation, learning strategies, cognitive resources, and self-regulation, whereas hopelessness may be detrimental.

As expected, relations for the positive deactivating emotion of relief and the negative activating emotions of anger, anxiety, and shame proved to be more complex. Specifically, besides negative relations to motivation, strategies, and cognitive resources, anxiety and shame correlated positively with self-reported use of rehearsal. This is in line with our theoretical expectations regarding the effects of emotions on flexible versus rigid learning strategies.

Test anxiety correlated significantly with academic achievement in both Studies 4 and 5. Again, test anxiety proved not to be the emotion most closely related to students' academic grades. Test-related joy, hope, and pride correlated significantly positively with achievement, whereas anger, anxiety, shame, and hopelessness correlated significantly negatively with achievement. Relations for the positive deactivating emotion of relief were near zero. The correlations for relief are probably based on ambivalent relations with achievement-correlated negative emotions, on the one hand, and with exam success triggering relief, on the other.

Relations to health complaints. As in Study 3, there were significant relations between negative emotions and a number of perceived health problems, including perceived cardiovascular problems, stomach problems, and sleep disturbances (Table VI).

Conclusions

The results of Studies 4 and 5 showed that the revised and final TEQ scales are reliable as well as internally and externally valid. Building on the first version of the TEQ, it proved possible to improve the reliability of the test hope scale, and to reduce the correlative association between the trait test anxiety and trait test hopelessness scales. Nevertheless, some of the trait scale intercorrelations of the TEQ were still relatively high. Specifically, this pertains to correlations within the two emotion groups of (a) joy, hope, and pride, and (b) anger, anxiety, and hopelessness. It turned out that the correlations within these groups could not be reduced to low values by using information about divergent item validities. They probably just represent the empirical reality. Joy, hope, and pride are inevitably related because hope and pride are specific pro- and retrospective variants of joy. Test-related anxiety, hopelessness, and shame also share many elements (e.g., failure-related worry cognitions). However, despite their close associations, emotions from these two groups may show differential patterns of relations to antecedents and effects. Furthermore, although closely related at the trait level, their intercorrelations at the level of state emotions are lower (Pekrun *et al.*, in prep.). Consequently, while it would seem possible to create combined scales measuring emotions of these two groups in integrative ways, such an integration would risk blurring empirically important distinctions.

As regards external validity, results were largely consistent with theoretical expectations. The relations of emotion scores to gender, perceived academic self-efficacy and control, motivation, learning strategies, irrelevant thinking, self-regulation, academic achievement, and perceived health problems indicate that test emotions may be of major relevance for learning, achievement, and health. However, as in Study 3, findings should be interpreted with caution because of their cross-sectional nature.

ENGLISH-LANGUAGE VERSION OF THE TEQ (STUDY 6)

To create versions of the scales for international use and cross-cultural comparisons, short versions of the scales were translated into English by a team of two Canadian and two German psychologists, two professional translators, and a number of German and Canadian graduate students (see Appendix for sample items in English). The translation was cross-checked by back-translation into German. The short versions of the scales used for translation were created by selecting items with favorable item statistics, taking care to ensure that all emotion components were covered. The empirical properties of the English version were analyzed in a North American student sample.

Method

The scales were administered to a sample of $N=389$ students enrolled in three undergraduate psychology courses at the University of Manitoba (234 females, 155 males; mean age 20.63 years, $SD=3.48$). Scale administration took place in the second to fourth weeks of the term. In addition to the emotion scales, the questionnaire included a short English version of the general self-esteem scale used in Study 3 (Schwarzer, 1986; six items, $\alpha=0.82$ in the present sample); Perry's scale on perceived

academic control (Perry, 1991; Perry *et al.*, 2001; 10 items, $\alpha = 0.83$); four scales from Pintrich's Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich *et al.*, 1991) measuring academic self-efficacy (five items, $\alpha = 0.82$), effort (six items, $\alpha = 0.61$), rehearsal (four items, $\alpha = 0.59$), and elaboration (six items, $\alpha = 0.73$); a short English version of the study interest scale used in Studies 3–5 (Winteler *et al.*, 1991; four items, $\alpha = 0.69$); a short English version of our scale on self-regulation of learning (four items, $\alpha = 0.72$); and grade point averages for the previous academic year (percentage scores; $M = 74.8\%$, $SD = 10.2\%$).

Results and Discussion

Score distributions, item statistics, and reliabilities. Item and scale statistics showed that the English versions of the emotion scales were characterized by sufficient variation of scores, as well as by good item-total correlations and reliabilities (Table II). Reliability coefficients were somewhat lower than for the long version of the German scales due to the reduction in item numbers.

Internal structural validity: Confirmatory factor analysis of scales. CFA was used to analyze internal structural scale validities (LISREL 8.53, Jöreskog and Sörbom, 2002). Adapting the procedure used by Hodapp and Benson (1997) to analyze the dimensionality of test anxiety, three structural models were competitively tested for each scale. In all three models, items were used as manifest variables. Model 1 (general factor model) was a one-factor model containing one latent emotion factor using all items of the scale as indicators. Model 2 (component factors model) contained four separate latent factors for the affective, cognitive, physiological, and motivational components of test emotions. Model 3 (hierarchical model) was basically identical with the component factors model, but additionally included a general factor representing the emotion in question as a second-order factor (see Fig. 1 for Models 2 and 3). Model 3 is nested within Model 2 and introduces additional restrictions, inevitably implying a reduction of empirical fit compared to Model 2.

Theoretically, our four-component conception of test emotions led us to expect (a) that the component factor models would show better fit than the one-factor models, and (b) that the hierarchical models would, at most, show a small loss of fit compared to the component factor models, implying that the component factors can legitimately be grouped under the umbrella of one emotion construct per scale. As shown by Table VII, both expectations were largely fulfilled. For all scales, the component factor models showed adequate fit to the data, and for six of the eight scales, the fit was clearly better than in the one-factor models. Furthermore, the hierarchical models did not imply a serious loss of fit, thus corroborating the component structure of our emotion constructs and the scales measuring them.

Convergent versus divergent item validities and intercorrelations of scales. The patterns of results for convergent versus divergent item validities (Table II) and correlations between scales (Table III) were similar to those for the German versions of the scales (see discussion above, Studies 4 and 5). Interestingly, within the group of the neighboring emotions of test anxiety, shame, and hopelessness, it was the relation between test shame and test hopelessness which was strongest, suggesting close links between these two emotions in the North American student sample.

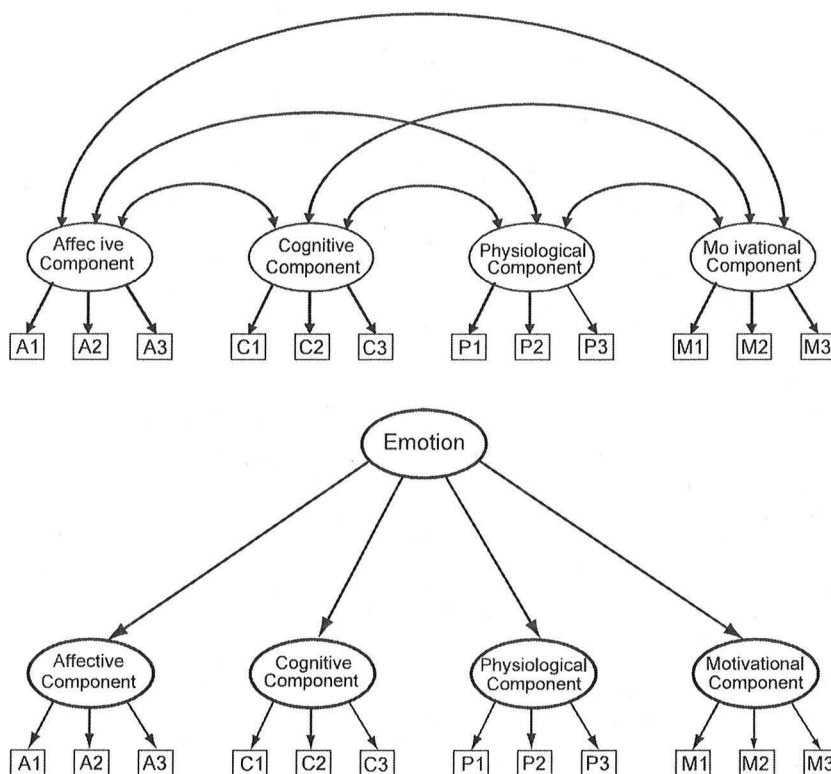


FIGURE 1 Structure of test emotions: Four-components model (upper part) and hierarchical model (lower part); items as manifest variables in each model.

External validity: Relations to gender, learning, and academic achievement. In line with the results for the German version of the TEQ, female students scored significantly higher on test anxiety than males. Gender differences for other test emotions were small (Table IV). The pattern of relations with variables of academic learning also converged with the findings for the German versions of the scales. A number of these correlations were even higher in the North American sample. Test joy, hope, and pride correlated positively, and hopelessness negatively, with self-efficacy, perceived academic control, effort, elaboration, and self-regulation. Again, negative correlations with these variables were weaker for test anxiety. In addition, test anxiety again showed a small positive correlation with rehearsal. Finally, there were a number of significant correlations with academic achievement. Similar to the German data sets, test anxiety correlated with achievement, but less so than other test emotions. Interestingly, the highest correlations with achievement were found for test pride and shame in the present sample. This corresponds with the importance attributed to these two emotions in achievement motivation theory (Weiner, 1985).

Conclusions

The English versions of the scales proved to be reliable, as well as valid in terms of internal factor structures and relations to demographic variables, learning, and

TABLE VII Summary of Confirmatory Factor Analysis of Test Emotion Scales

	<i>Model</i>	<i>Fit index</i>					
		χ^2	<i>df</i>	χ^2/df	<i>GFI</i>	<i>CFI</i>	<i>RMSEA</i>
Joy	1	119	34	3.50	0.94	0.89	0.081
	2	64	28	2.29	0.97	0.95	0.058
	3	71	30	2.37	0.96	0.94	0.060
Hope	1	75	20	3.75	0.95	0.94	0.085
	2	43	17	2.53	0.97	0.97	0.063
	3	43	17	2.53	0.97	0.97	0.063
Pride	1	113	32	3.53	0.94	0.94	0.082
	2	64	26	2.46	0.97	0.97	0.062
	3	86	28	3.07	0.96	0.96	0.074
Relief	1	39	13	3.00	0.97	0.96	0.072
	2	26	12	2.17	0.98	0.98	0.056
	3	25	12	2.08	0.98	0.98	0.056
Anger	1	97	33	2.94	0.95	0.95	0.071
	2	45	27	1.67	0.98	0.98	0.042
	3	52	29	1.79	0.97	0.98	0.046
Anxiety	1	228	53	4.30	0.91	0.92	0.093
	2	95	47	2.02	0.96	0.98	0.052
	3	110	49	2.24	0.95	0.97	0.057
Shame	1	107	33	3.24	0.95	0.95	0.076
	2	100	27	3.70	0.95	0.95	0.084
	3	104	29	3.59	0.95	0.95	0.082
Hopelessness	1	92	44	2.09	0.96	0.98	0.054
	2	82	38	2.16	0.96	0.98	0.055
	3	83	40	2.08	0.96	0.98	0.053

Note. Models 1/2/3: General factor model, four-components factor model, hierarchical model.

academic achievement. The internal and external scale characteristics were similar to those of the original German versions of the scales. Any differences between findings for the North American sample and the German samples were in the same range as the variation between the different German samples (Studies 3–5), lending support to the hypothesis of cross-cultural equivalence of the scales. In sum, it seems that the English version of the TEQ is as ready to be used in research on test emotions as the German version is.

GENERAL DISCUSSION

Over the past 65 years, measurement and scientific analysis of test emotions has focused on test-related anxiety. Today, a host of test anxiety inventories are available, whereas instruments measuring test emotions other than anxiety are lacking. Advances in the study of different test emotions, however, may be just as dependent on the development of suitable measurement instruments as advances in test anxiety research have been over the past decades.

In the research presented here, a questionnaire was developed which measures test-related joy, hope, pride, relief, anger, anxiety, shame, and hopelessness. The strategy of developing scales implied to develop domain-related theory and conduct exploratory

analysis first, before constructing and validating items and scales. In the theory phase, we defined "test emotions" as emotions subjectively relating to taking tests or exams, and developed a four-components conception specifying affective, cognitive, physiological, and motivational components for a number of major test emotions. Assumptions on external relations of test emotions were derived from Pekrun's control-value theory of achievement emotions (Pekrun, 2000; Pekrun *et al.*, 2002a, b). This theory implies that test emotions may depend on achievement-related appraisals of control and values, influence learning processes and performance, and are linked to resulting achievement by feedback loops of reciprocal causation over time.

The findings of our exploratory analyses confirm that anxiety is a key emotional experience in test and exam situations, but they are also in line with assuming that anxiety is not the only human emotion relevant in such situations. Positive emotions like hope, relief and joy, and negative emotions like anger, shame, and hopelessness may be important as well, according to this analysis. Findings thus suggest that students' exam-related emotional life is richer than prevailing views on the importance of test anxiety seem to suggest. Any account of emotions relating to tests and exams which narrows the range of feelings to anxiety only may be at risk of misrepresenting reality.

Based on these findings and on a four-components conception of test emotions differentiating affective, cognitive, physiological, and motivational elements of emotions, the scales of the Test Emotions Questionnaire were developed and fine-tuned in German before being translated into English. Findings on the scales demonstrate that scale scores show sufficient variation, are reliable, and are structurally valid in terms of representing the proposed component structure of test emotions. One specific finding of relevance to test anxiety research is that the subscales of the Reactions to Tests questionnaire (Sarason, 1984) used to operationalize test anxiety in the first version of the TEQ were strongly correlated with test hopelessness (Study 3). To the extent that this finding is generalizable across subject samples and test anxiety inventories, it would imply that contemporary measures of test anxiety might be at risk of lacking divergent validity. Analyzing these measures in conjunction with measures of other test emotions may help to determine to what extent this is the case, and how test anxiety measurement can be improved such that sufficient discrimination of test anxiety from neighboring emotions is achieved.

However, there may be limits to reducing correlations between trait test emotion measures. High interindividual covariation within groups of neighboring trait test emotions may simply be the empirical reality (see also Watson and Clark, 1992). Specifically, judging from our findings on the TEQ scales, this may be true for the two emotion groups of test-related joy, hope, and pride, and of test-related anxiety, shame, and hopelessness. Concerning the first group, hope and pride may be regarded as cognitively specific variants of the primary emotion of joy, as argued above. Hope can be interpreted as anticipatory joy, implying uncertainty about future positive outcomes, and pride as concurrent or retrospective joy, implying internal causal attributions of positive outcomes (see Weiner, 1985, for cognitive appraisals implied by pride). Consequently, the constructs of test-related joy, hope, and pride share emotional contents, probably also share situational antecedents, and should therefore correlate empirically. Concerning the second group, anxiety, shame, and hopelessness refer to different primary emotions. Nevertheless, they may also share components and

antecedents. Specifically, worries may be typical of all three emotions, implying that they should also correlate.

In spite of the high correlations within these two groups of trait test emotions, it may pay to measure them separately, for at least two reasons. First, there seems to be differentiation at the level of state emotions (Pekrun *et al.*, in prep.). Differentiating test emotions at the trait level as well ensures that state and trait level measures of test emotions are conceptually equivalent. Second, in spite of content overlap and partially shared antecedents, there may also be emotion-specific, differential antecedents and effects, even at the level of trait emotions. For example, whereas motivational effects of the activating emotions of anxiety and shame may be ambivalent, hopelessness may clearly be detrimental. Consequently, the negative correlations with academic interest and effort found for all three emotions in the present research (Studies 3–6) were consistently higher for hopelessness than for anxiety and shame. Another, specific example would be suicidal ideation triggered by exam-related hopelessness rather than just by anxiety, as reported by one participant in our exploratory studies. Concerning differential antecedents, attributional theories and control-value theory imply that joy, hope and pride, as well as anxiety, shame and hopelessness, should be linked to appraisals pertaining to causal dimensions and levels of subjective certainty in differential ways, although it may be difficult to detect these differences by trait-level self-report methods (Weiner, 1985; Pekrun, 2000).

The external relations of the TEQ scales to demographic variables, personality, learning, academic achievement, and perceived health proved to be largely congruent to theoretical expectations. In line with previous research, our findings implied that test anxiety is correlated with outcome variables. However, they also demonstrated that test emotions other than anxiety can show even stronger relations to variables of academic learning and achievement. This is in line with the assumptions of Pekrun's control-value theory (Pekrun *et al.*, 2002a, b) and other theories on the ambivalent nature of test anxiety effects (see Zeidner, 1998), and implies that test anxiety is neither the only relevant test emotion, nor necessarily the most important one.

Test joy, hope, and pride correlated positively with academic self-efficacy, interest and effort, use of learning strategies, self-regulation of learning, and academic achievement. The correlations for relief were weak and less consistent. This may be due to ambivalent effects of positive deactivating emotions like relief, and to the close relation between relief and preceding negative emotions. For anger, anxiety, shame, and hopelessness, correlations were negative. However, as expected, anxiety tended to correlate positively with the rigid learning strategy of rehearsal. Most of the correlations between test emotions and variables of learning and achievement were consistent across four independent empirical studies and German versus North American student samples, implying cumulative evidence corroborating the robustness of linkages between these emotions and students' performance.

There are some limitations to the present research which should be addressed in future studies. First, the properties of the empirical scales need to be confirmed in different populations and samples, beyond the samples of university students addressed in the present series of studies. The English version of the TEQ was analyzed in a sample of North American students; it remains to be seen whether the empirical qualities of its scales are stable across other English-speaking samples. Second, where reliability is concerned, we gathered data on internal consistency, but the temporal

stability (retest reliability) of the TEQ's trait test emotion scales is yet to be examined. Finally, our findings on the external relations of test emotions should be interpreted with caution because of their cross-sectional status. For example, it can be assumed that the correlations reported for test emotions and academic achievement only partially reflect effects of test emotions on achievement. As implied by the control-value theory, reverse effects of students' success and failure in exams on the development of test emotions may be equally important (for reciprocal causation of test anxiety and achievement, see Meece *et al.*, 1990; Pekrun, 1992a; Schnabel, 1998). Future studies should analyze the usability of the TEQ in diverse age groups, institutions, and countries, as well as its retest reliability and predictive validity in explaining long-term individual development and outcomes.

Author Note

This research was supported by two research grants to the first author from the German Research Foundation (Deutsche Forschungsgemeinschaft, Pe 320/7-1, 320/7-3) and by a grant to the first and third authors from the German American Academic Council.

References

- Brown, C.H. (1938). Emotional reactions before examinations: II. Results of a questionnaire. *Journal of Psychology*, *5*, 11–26.
- Brown, C.H. and van Gelder, D. (1938). Emotional reactions before examinations: I. Physiological changes. *Journal of Psychology*, *5*, 1–9.
- Elliot, A.J. and McGregor, H.A. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *76*, 628–644.
- Elliot, A.J. and McGregor, H.A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*, 501–519.
- Eysenck, M. and Calvo, M.G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion*, *6*, 409–434.
- Feldman Barrett, L. and Russell, J.A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*, 967–984.
- Heckhausen, H. (1991). *Motivation and action*. Springer, Berlin.
- Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, *58*, 47–77.
- Hodapp, V. and Benson, J. (1997). The multidimensionality of test anxiety: A test of different models. *Anxiety, Stress and Coping*, *10*, 219–244.
- Houston, D.M. (1995). Vulnerability to depressive mood reactions: Retesting the hopelessness model of depression. *British Journal of Social Psychology*, *34*, 293–302.
- Jerusalem, M. and Schwarzer, R. (1986). Fragebogen zur Erfassung von "Selbstwirksamkeit" [Questionnaire for assessing self efficacy]. In: Schwarzer, R. (Ed.), *Skalen zur Befindlichkeit und Persönlichkeit*, pp. 15–28. Free University of Berlin, Institute for Psychology, Berlin.
- Joreskog, K. and Sorbom, D. (2002). *LISREL 8.53* (Update December 2003). Scientific Software International, Chicago.
- Liebert, R.M. and Morris, L.W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, *20*, 975–978.
- Mandler, G. and Sarason, S.B. (1952). A study of anxiety and learning. *Journal of Abnormal and Social Psychology*, *47*, 166–173.
- Meece, J.L., Wigfield, A. and Eccles, J.S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, *82*, 60–70.
- Meinhardt, J. and Pekrun, R. (2003). Attentional resource allocation to emotional events: An ERP study. *Cognition and Emotion*, *17*, 477–450.
- Metalsky, G.I., Halberstadt, L.J. and Abramson, L.Y. (1987). Vulnerability to depressive mood reactions: Toward a more powerful test of the diathesis stress and causal mediation components of the reformulated theory of depression. *Journal of Personality and Social Psychology*, *52*, 386–393.

- Molfenter, S. (1999). *Prüfungsempfindungen bei Studierenden* [Test emotions in university students]. Unpublished dissertation. Institute of Psychology, University of Regensburg.
- Nicholls, J.G. (1976). When a scale measures more than it denotes: The case of the Test Anxiety Scale for Children. *Journal of Consulting and Clinical Psychology*, **44**, 976–985.
- Pekrun, R. (1992a). Expectancy value theory of anxiety: Overview and implications. In: Forgas, D.G., Sosnowski, T. and Wrzesniewski, K. (Eds.), *Anxiety: Recent developments in self appraisal, psychophysiological and health research*, pp. 23–41. Hemisphere, Washington, DC.
- Pekrun, R. (1992b). Kognition und Emotion in studienbezogenen Lern und Leistungssituationen: Explorative Analysen [Cognition and emotion in academic situations of learning and achievement: Exploratory analysis]. *Unterrichtswissenschaft*, **20**, 308–324.
- Pekrun, R. (2000). A social cognitive, control value theory of achievement emotions. In: Heckhausen, J. (Ed.), *Motivational psychology of human development*, pp. 143–163. Elsevier, Oxford, UK.
- Pekrun, R. and Frese, M. (1992). Emotions in work and achievement. In: Cooper, C.L. and Robertson, I.T. (Eds.), *International Review of Industrial and Organizational Psychology*, **7**, 153–200.
- Pekrun, R., Goetz, T., Spangler, G. and Perry, R.P. (in prep.). *Beyond test anxiety: The impact of positive and negative test emotions on students' exam performance*. Department of Psychology, University of Munich.
- Pekrun, R., Goetz, T., Titz, W. and Perry, R.P. (2002a). Academic emotions in students' self regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist*, **37**, 91–106.
- Pekrun, R., Goetz, T., Titz, W. and Perry, R.P. (2002b). Positive emotions in education. In: Frydenberg, E. (Ed.), *Beyond coping: Meeting goals, visions, and challenges*, pp. 149–174. Elsevier, Oxford, UK.
- Perry, R.P. (1991). Perceived control in college students: Implications for instruction in higher education. In: Smart, J. (Ed.), *Higher education: Handbook of theory and research*, Vol. 7, pp. 1–56. Agathon, New York.
- Perry, R.P., Hladkyj, S., Pekrun, R. and Pelletier, S. (2001). Academic control and action control in college students: A longitudinal field study. *Journal of Educational Psychology*, **93**, 776–789.
- Pintrich, P.R., Smith, D.A.F., Garcia, T. and McKeachie, W. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. University of Michigan, Ann Arbor, MI.
- Quast, H.H., Jerusalem, M. and Sarason, I.G. (1986). Reaktionstendenzen gegenüber Tests RTT G (Reactions to Tests). In: Schwarzer, R. (Ed.), *Skalen zur Befindlichkeit und Persönlichkeit*, pp. 139–155. Institute for Psychology, Free University of Berlin.
- Sarason, I.G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology*, **44**, 929–938.
- Scherer, K.R. (1984). On the nature and function of emotion: A component process approach. In: Scherer, K.R. and Ekman, P. (Eds.), *Approaches to emotion*, pp. 293–317. Erlbaum, Hillsdale, NJ.
- Scherer, K.R., Wallbott, H.G. and Summerfield, A.B. (1986). *Experiencing emotion: A cross cultural study*. Cambridge University Press, Cambridge, UK.
- Schiefele, U., Krapp, A., Wild, K. P. and Winteler, A. (1993). Der "Fragebogen zum Studieninteresse", FSI [The Questionnaire on Study Interest FSI]. *Diagnostica*, **39**, 335–351.
- Schnabel, K. (1998). *Prüfungsangst und schulisches Lernen* [Test anxiety and learning at school]. Waxmann, Muenster.
- Schwarzer, R. (Ed.). (1986). *Skalen zur Befindlichkeit und Persönlichkeit* [Scales on psychological health and personality]. Institute for Psychology, Free University of Berlin.
- Schutz, P. and Davis, H.A. (2000). Emotions and self regulation during test taking. *Educational Psychologist*, **35**, 243–256.
- Smith, C.A. and Ellsworth, P.C. (1987). Patterns of appraisal and emotion related to taking an exam. *Journal of Personality and Social Psychology*, **52**, 475–488.
- Spangler, G., Pekrun, R., Kramer, K. and Hofmann, H. (2002). Students' emotions, physiological reactions, and coping in academic exams. *Anxiety, Stress and Coping*, **15**, 383–340.
- Spielberger, C.D. (1980). *Test Anxiety Inventory: Preliminary professional manual*. Consulting Psychologist Press, Palo Alto, CA.
- Spielberger, C.D., Anton, W.D. and Bedell, J. (1976). The nature and treatment of test anxiety. In: Zuckerman, M. and Spielberger, C.D. (Eds.), *Emotions and anxiety: New concepts, methods, and applications*, pp. 317–344. Erlbaum, Hillsdale, NJ.
- Stengel, E. (1936). Prüfungsangst und Prüfungsneurose [Test anxiety and test neurosis]. *Zeitschrift für psychoanalytische Pädagogik*, **10**, 300–320.
- Titz, W. (2001). *Emotionen von Studierenden in Lernsituationen* [Students' emotions in situations of learning]. Waxmann, Muenster.
- Watson, D. and Clark, L.A. (1992). Affects separable and inseparable: On the hierarchical arrangement of the negative affects. *Journal of Personality and Social Psychology*, **62**, 489–505.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, **92**, 548–573.

- Wild, K. P. and Schiefele, A. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens [Learning strategies at university: Findings on the factorial structure and reliability of a new questionnaire]. *Zeitschrift fuer Differentielle und Diagnostische Psychologie*, 15, 185–120.
- Winteler, A., Schiefele, U., Krapp, A. and Wild, K. P. (1991). *Skalen zu Studieninteresse und Lernstrategien im Studium* [Scales on study interest and learning strategies]. University of the German Army, Neubiberg/Munich, Germany.
- Zeidner, M. (1998). *Test anxiety: the state of the art*. Plenum, New York.
- Zerssen, D. (1976). *Die Beschwerdenliste* [Inventory of health complaints]. Beltz, Weinheim.

APPENDIX

Test Emotions Questionnaire (TEQ): scales and sample items

TEST JOY

- I look forward to the exam (b)
 Before taking the exam, I sense a feeling of eagerness (b)
 For me the test is a challenge that is enjoyable (d)

TEST HOPE

- I am optimistic that everything will work out fine (b)
 I have great hope that my abilities will be sufficient (b)
 I am very confident (d)

TEST PRIDE

- I am proud of myself (a)
 To think about my success makes me feel proud (a)
 After the exam I feel ten feet taller because I'm so proud (a)

TEST RELIEF

- I feel freed (a)
 I feel very relieved (a)
 I finally can breathe easy again (a)

TEST ANGER

- I get angry (d)
 I am fairly annoyed (a)
 I get so angry, I start feeling hot and flushed (a)

TEST ANXIETY

- I feel panicky when writing an exam (d)
 I worry whether I will pass the exam (d)
 I get so nervous I wish I could just skip the exam (b)

TEST SHAME (10)

- I feel ashamed (a)
 My marks embarrass me (a)
 I am ashamed of my poor preparation (d)

TEST HOPELESSNESS

I feel hopeless (d)

I have lost all hope that I have the ability to do well on the exam (d)

My hopelessness robs me of all my energy (b)

Note. b/d/a =before, during, after the exam.