

The AEQ-S: A short version of the Achievement Emotions Questionnaire

Maik Bieleke^{a,*}, Katarzyna Gogol^{b,c}, Thomas Goetz^a, Lia Daniels^d, Reinhard Pekrun^{e,f,g}^a Department of Developmental and Educational Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria^b Empirical Educational Research, University of Konstanz, Konstanz, Germany^c Empirical Educational Research, Thurgau University of Teacher Education, Kreuzlingen, Switzerland^d Department of Educational Psychology, University of Alberta, Edmonton, Canada^e Department of Psychology, University of Essex, Colchester, United Kingdom^f Institute for Positive Psychology and Education, Australian Catholic University, Sydney, Australia^g Department of Psychology, University of Munich, Munich, Germany

ARTICLE INFO

Keywords:

Achievement Emotions Questionnaire (AEQ)

Achievement emotions

Control-value theory

Short scale

Assessment

ABSTRACT

The Achievement Emotions Questionnaire (AEQ) is a well-established instrument for measuring achievement emotions in educational research and beyond. Its popularity rests on the coverage of the component structure of various achievement emotions across different academic settings. However, this broad conceptual scope requires the administration of 6 to 12 items per scale ($Mdn = 10$), which limits the applicability of the AEQ in empirical studies that necessitate brief administration times. We therefore developed the AEQ-S, a short version of the AEQ, with only 4 items per scale that nevertheless maintain the conceptual scope of the instrument. We validated the AEQ-S based on a reanalysis of Pekrun, Goetz, Frenzel, Barchfeld, and Perry's (2011) dataset ($N = 389$ university students) and by administering them to a new and independent validation sample ($N = 471$ university students). Despite their brevity, the AEQ-S scales achieved satisfactory reliability and correlated substantially with the original AEQ scales. Moreover, structural relationships and intercorrelations between the scales and their relations with external measures of antecedents and outcomes of achievement emotions were highly similar for the AEQ-S and AEQ scales. These findings suggest that the AEQ-S is a suitable substitute for the AEQ when administration time is limited.

1. Introduction

Academic settings trigger a broad variety of different emotions – students enjoy to learn something new, are proud of their test results, get annoyed at the amount of their homework, or feel bored during a dull class. Until about 20 years ago, educational researchers studied only a few of these achievement emotions, predominantly as single emotions and within specific settings only (e.g., test anxiety; Zeidner, 1998). One major reason for this research gap was the lack of comprehensive instruments for measuring achievement emotions. This lack was addressed with the introduction of the *Achievement Emotions Questionnaire* (AEQ; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011; Pekrun, Goetz, Titz, & Perry, 2002). The AEQ is a theory-grounded measurement instrument that allows the assessment of various achievement emotions across different settings. Since its introduction, it has become a popular and well-established instrument in educational psychology and beyond, as reflected in more than 3,700 Google Scholar citations (as of January

2021) of its initial publication (Pekrun et al., 2002) and about 1,400 citations of its current version (Pekrun et al., 2011).

The AEQ laid the foundations for pioneering research on emotional experiences in various educational and achievement settings. One key to this success is the conceptual scope of the instrument, which allows a comprehensive analysis of achievement emotions. The AEQ comprises 24 scales assessing the nine trait emotions of achievement-related enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, and boredom. These scales are organized in three sections pertaining to class-, learning-, and test-related academic settings. The items cover emotional experiences either before, during, or after the corresponding setting, and they measure the affective, cognitive, motivational, and physiological components of each emotion. This broad scope allows researchers to study a comprehensive range of emotional experiences in academic settings. The downside of this breadth of coverage, however, is the resulting length of the questionnaire. The AEQ contains a total of 232 items, with 80 items pertaining to class-related emotions, 75 items

* Corresponding author at: University of Vienna, Universitaetsstrasse 7, 1010 Vienna, Austria.

E-mail address: maik.bieleke@univie.ac.at (M. Bieleke).

<https://doi.org/10.1016/j.cedpsych.2020.101940>

to learning-related emotions, and 77 items to test-related emotions. This length can make it difficult or even impossible to administer the AEQ in empirical studies in educational research because testing time is often limited (e.g., large-scale assessments, longitudinal studies with measurement burst designs), space for questions might be restricted or expensive, and long questionnaires might alter the experience they measure. This problem also pertains to situations in which only selected scales are administered rather than the entire AEQ (e.g., with a focus on specific emotions in a particular academic setting) as the AEQ scales each comprise between 6 and 12 items with a median number of 10 items.

As a consequence, it is plausible that few researchers use the AEQ in full length but instead use select scales only or shorten the scales for their purposes. To confirm this presumption, we analyzed the length of the scales reported in 853 papers citing the most recent publication of the AEQ (Pekrun et al., 2011) as of March 2019. Fig. 1 shows the distribution of the observed scale lengths. We found a median scale length of 5 items, meaning that about 50% of the scale versions used in the literature are shorter than even the shortest original scales of the AEQ. Moreover, the distribution is multimodal and its shape suggests three main fields of application and their corresponding requirements regarding scale length. A first peak appears for scales with single items and suggests a need for ultra-short assessments (e.g., in experience sampling studies). A second peak appears for scales with 3 to 6 items and suggests a need for assessments that balance brevity and reliability (e.g., as part of a battery of questionnaires). Finally, a third peak appears for scales with 8 to 11 items and suggests the need for comprehensive assessments with maximum reliability (e.g., when the main interest is on measuring one specific achievement emotion). Taken together, these findings suggest a substantial interest in as well as an active use of short AEQ scales in the scientific community. Unfortunately, however, these short scales have so far been compiled in an *ad hoc* manner, typically without proper analysis of their psychometric quality. The specific items are not always fully reported and they are likely to differ widely across

studies, which makes it difficult to compare results and to adapt any of the different variants to new research questions.

Another indication that there is a demand for a short version of the AEQ comes from published questionnaires that were developed on the basis of the AEQ. All of these adaptations use scales with fewer items than the full AEQ scales. For instance, the AEQ-M (Pekrun, Goetz, & Frenzel, 2005; Pekrun, Lichtenfeld, Marsh, Murayama, & Goetz, 2017; Pekrun, Murayama, Marsh, Goetz, & Frenzel, 2019) assesses students' achievement emotions in mathematics with a median scale length of 8 items. Even shorter are a version for use with elementary school students that contains 2 to 5 items per scale (AEQ-ES; Lichtenfeld, Pekrun, Stupnisky, Reiss, & Murayama, 2012) and a version validated for use with pre-adolescents that comprises 4 items per scale (AEQ-PA; Peixoto, Mata, Monteiro, Sanches, & Pekrun, 2015). Similarly, adaptations to other cultures and languages such as the AEQ for Filipinos (King, 2010) and for Italians (Raccanello, 2015) use scales with only 2 or 3 items per scale, respectively. While these examples highlight researchers' interest in developing short scales for measuring achievement emotions, none of these adaptations achieves the same scope as the original version (e.g., with regard to number of emotions and academic settings, or systematic coverage of emotion components), implying that they are of limited value compared with the original AEQ.

1.1. The present research

The goal of the present study was, therefore, to thoroughly develop a short version of the AEQ and to scrutinize its psychometric properties by conducting systematic comparisons with the original AEQ (Pekrun et al., 2011). To achieve this aim, we first used data on the AEQ scales developed by Pekrun et al. (2011) to construct a short version of the AEQ, the AEQ-S. This allowed us to compare the psychometric properties of the AEQ and the AEQ-S in this dataset. We then administered the AEQ-S along with all other scales used by Pekrun et al. (2011) to an independent validation sample and examined its psychometric

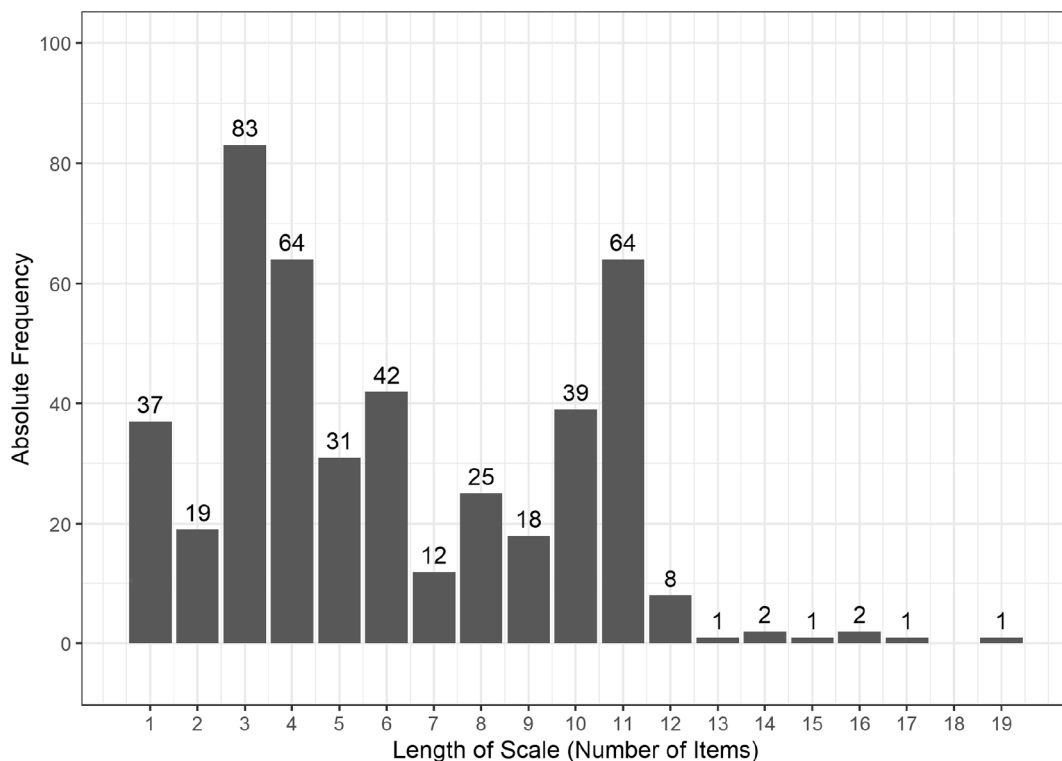


Fig. 1. Histogram showing the length of AEQ scales used in the literature. The observed median length was 5 items, which is half the median length of the full AEQ scales (10 items). Note that many studies used more than one scale.

properties in comparison to the AEQ.

2. Methods

2.1. Samples

The analyses of the present study are based on two independent datasets. First, we reanalyzed the data described by Pekrun et al. (2011), which comprised a sample of 389 undergraduate students (234 females, 60.2%) with a mean age of $M = 20.63$ years ($SD = 3.48$) at a large, Midwestern Canadian university. Most students were enrolled in study programs in arts (42.7%), followed by management (15.4%), science (12.7%), and nursing (8.1%). Second, we conducted a validation study for which we collected new data from an independent sample of 471 undergraduate students at another Canadian university (171 females, 36.3%; 7 participants did not indicate their gender, 1.48%) with a mean age of $M = 19.73$ years ($SD = 1.81$; 4 participants did not indicate their age). Most students in the validation study were enrolled in science and engineering programs (54.1%), followed by arts (28.0%), business (7.9%), and education (6.8%). They were recruited from the university's participant pool, in which approximately 3,500 undergraduate students can take part in research for course credit each semester. Everyone in this pool would have had the chance to sign up for the study, although we set an upper limit of 500 participants. In both studies, participants completed the questionnaire in a paper-and-pencil format.

2.2. Measures

The measures reported in Pekrun et al.'s (2011) study and those used in the validation study were identical, except that the validation study used the AEQ-S scales rather than the full AEQ scales. The same measures were chosen to facilitate direct comparisons between the AEQ-S and the AEQ, which were the main focus of our study. In both studies, students provided answers to items on 5-point Likert scales ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). We averaged the item scores to obtain composite scores and handled missing values with full information maximum likelihood (FIML) estimation rather than listwise deletion as in Pekrun et al. (2011). This led to some negligible deviations of our results for the full AEQ from those reported by Pekrun et al. (2011).

2.2.1. Achievement Emotions Questionnaire (AEQ)

Pekrun et al. (2011) introduced the AEQ with 24 scales and a total of 232 items (6 to 12 items per scale) which measure (a) enjoyment, hope, pride, anger, anxiety, shame, hopelessness, and boredom in class-related settings, (b) the same emotions in learning-related settings, and (c) enjoyment, hope, pride, relief, anger, anxiety, shame, and hopelessness in test-related settings. In our reanalysis of these authors' data, we selected four items for each scale which form the AEQ-S with a total of 96 items. As such, in the reanalysis the AEQ-S scales were part of the full AEQ scales. In contrast, in our validation study we only used the 96 items of the AEQ-S.

2.2.2. Selection of the AEQ-S items

To construct the scales of the AEQ-S, we selected items so as to cover all four components of each emotion considered in the AEQ (i.e., affective, cognitive, motivational, and physiological). We determined unidimensional factor models for each construct (e.g., learning-related anxiety). We then used these models to identify the four items with the highest factor loadings among the items representing the affective, the cognitive, the motivational, and the physiological component, respectively. Because the affective component represents the necessary, core constituent of emotions (e.g., Pekrun, Muis, Frenzel, & Goetz, 2017), we considered the conceptual scope of the affective items as an additional selection criterion. As the AEQ does not assess the physiological component of hope, we instead used one item from one of the other components (see Appendix for the final selection of items).

2.2.3. Antecedents of academic emotions

As antecedents of academic emotions, perceived academic control, self-efficacy, and task value were measured by Pekrun et al. (2011) and in the validation study.

Academic control and self-efficacy. Achievement-related subjective academic control and self-efficacy were measured with a 10-item version of the *Perceived Academic Control Scale* (Perry, Hladkyj, Pekrun, & Pelletier, 2001) and a 5-item version of the *Self-efficacy for Learning and Performance Scale* of the *Motivated Strategies for Learning Questionnaire* (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1991), respectively. The Perceived Academic Control Scale measures students' perceived influence on their academic performance (e.g., "I have a great deal of control over my academic performance in my courses", "The more effort I put in my courses, the better I do in them"). The Self-efficacy for Learning and Performance Scale measures how confident students are of mastering academic tasks and getting good grades (e.g., "I'm confident I can do an excellent job on the assignments and tests in courses at university"). Cronbach's α for academic control and self-efficacy, respectively, was .83 and .82 in Pekrun et al.'s (2011) study and .80 and .88 in the validation study.

Task value. Perceived task value was assessed with a 4-item version of *Task Value Scale* of the MSLQ (Pintrich et al., 1991), which measures students' perceptions of how interesting, important, and useful academic tasks are (e.g., "Understanding the subject matter of courses at university is very important to me", "I am very interested in the content areas of courses at university"). Cronbach's α was .69 in Pekrun et al.'s (2011) study and .80 in the validation study.

2.2.4. Consequences of achievement emotions

As consequences of academic emotions, motivation, learning strategies, and self-regulation versus external regulation of learning were measured in Pekrun et al.'s (2011) study and in the validation study.

Motivation: Intrinsic goal orientation, extrinsic goal orientation, and effort regulation. We measured motivation with the *Intrinsic Goal Orientation Scale*, the *Extrinsic Goal Orientation Scale*, and the *Effort Regulation Scale* of the MSLQ (Pintrich et al., 1991), each of which comprises 4 items. The Intrinsic Goal Orientation Scale captures whether students are motivated by interest and curiosity (e.g., "In classes at university, I prefer course material that arouses my curiosity, even if it is difficult to learn"), while the Extrinsic Goal Orientation Scale assesses motivation based on grades and performance comparisons (e.g., "Getting good grades in classes at university is the most satisfying thing for me right now"). The Effort Regulation Scale captures how much effort students put into learning (e.g., "I work hard to do well in my classes even if I don't like what we are doing"). Cronbach's α was .51, .68, and .61, respectively, for intrinsic motivation, extrinsic motivation, and effort regulation in Pekrun et al.'s (2011) study and .64, .69, and .71 in the validation study.

Learning strategies, self-regulation of learning, and external regulation of learning. We assessed learning strategies with the 6-item *Elaboration Scale* and the 4-item *Rehearsal Scale* of the MSLQ (Pintrich et al., 1991). The Rehearsal Scale measures whether students recite and name items to learn them ("When studying for my classes, I practice saying the material to myself over and over"), while the Elaboration Scale measures the use of strategies like paraphrasing and creating analogies (e.g., "When reading for my classes, I try to relate the material to what I already know"). Cronbach's α of the elaboration and rehearsal scales was $\alpha = .73$ and .59, respectively, in Pekrun et al.'s (2011) study and .73 and .56 in the validation study.

We used the 4-item *Perceived Self-Regulation of Learning Scale* and the 4-item *Perceived External Regulation of Learning Scale* (Goetz, 2004) to measure students' self-regulated learning (e.g., "When studying, I set my own goals that I want to attain", "When studying difficult material, I decide for myself which strategy to use", "I am able to evaluate for myself how I make progress at learning") and externally regulated learning (e.g., "The way I study largely depends on the professor's

recommendations”, “When studying, I entirely rely on the readings I am given”). Cronbach’s α for self-regulation and external regulation, respectively, was .72 and .57 in Pekrun et al.’s (2011) study and .77 and .54 in the validation study.

2.2.5. Academic performance

We operationalized student’s performance as their self-reported grade point average attained over the academic year prior to the study.

2.3. Statistical analysis

We focus on examining the psychometric qualities of the AEQ-S. The psychometric qualities of the AEQ have been described in detail by Pekrun et al. (2011) and are reported here only to the extent necessary for evaluating the AEQ-S.

According to Gogol and colleagues (2014; see also Marsh, Ellis, Parada, Richards, & Heubeck, 2005; Smith, McCarthy, & Anderson, 2000), short scales that are derived from full scales should meet three crucial requirements: They have to reliably measure the constructs under investigation, they must reproduce the information obtained from the full scales, and they have to be embedded in the same nomological network as the full scales. To assess *reliability*, we compared Cronbach’s α of the AEQ-S scales with Cronbach’s α of the AEQ scales as reported by Pekrun et al. (2011). To assess *information reproduction*, we computed correlations between the AEQ-S and the AEQ scales based on Pekrun et al.’s (2011) data. Because the AEQ-S scales were embedded in the AEQ scales in this analysis, these correlations might be inflated due to common method variance. We therefore additionally computed corrected correlation coefficients using a formula suggested by Levy (1967), which removes the covariance between short and full scales resulting from correlated measurement error (Girard & Christensen, 2008). Finally, to compare the *nomological networks* of the AEQ and the AEQ-S scales, we employed confirmatory factor analysis (CFA) to

examine whether the AEQ-S reproduces the two-facet structure of achievement emotions (i.e., nine different emotions nested within three different achievement settings) that was found to best describe data collected with the AEQ (Pekrun et al., 2011). Moreover, we determined correlations of emotions within settings (e.g., class-related enjoyment vs. hope) and correlations of settings within emotions (e.g., class-related vs. learning-related enjoyment). Finally, we determined the correlations of achievement emotions as assessed by the short and full AEQ scales with external measures of the proposed antecedents and effects (see Pekrun et al., 2011).

3. Results

3.1. Reliability

We first determined scale statistics and computed the reliability of the AEQ-S and the AEQ scales in terms of Cronbach’s α (see Tables 1 and 2). The reliabilities of the AEQ scales ranged from $\alpha = .75$ to $.93$ ($M_\alpha = .85$). The reliabilities of the AEQ-S scales were smaller but still satisfactory, ranging from $\alpha = .64$ to $.88$ ($M_\alpha = .76$) in the reanalysis of the Pekrun et al. (2011) dataset and from $\alpha = .63$ to $.84$ ($M_\alpha = .76$) in the validation study. Thus, Cronbach’s α of the AEQ-S scales was on average $M_\Delta = .09$ smaller than Cronbach’s α of the AEQ scales, both in the reanalysis and in the validation study. This finding was robust across class-related, learning-related, and test-related emotions ($M_\Delta = .09/.09/.09$ in the reanalysis and $M_\Delta = .09/.09/.08$ in the validation study).

3.2. Information reproduction

Table 2 shows substantial relations between the AEQ scales and the AEQ-S scales in the reanalysis for class-related ($r = .88$ to $.95$), learning-related ($r = .89$ to $.95$), and test-related emotions ($r = .88$ to $.96$). The average correlation across emotions and settings was $M_r = .92$ (using

Table 1
Scale statistics of the AEQ scales and the AEQ-S scales.

	AEQ				AEQ-S							
	<i>M</i>	<i>SD</i>	Skewness	$\bar{r}_{i(t-i)}$	Reanalysis				Validation			
					<i>M</i>	<i>SD</i>	Skewness	$\bar{r}_{i(t-i)}$	<i>M</i>	<i>SD</i>	Skewness	$\bar{r}_{i(t-i)}$
<i>Class Settings</i>												
Enjoyment	3.19	0.65	−0.12	.56	3.25	0.75	−0.14	.55	3.07	0.77	0.05	.53
Hope	3.42	0.58	−0.19	.49	3.50	0.67	−0.36	.52	3.27	0.79	−0.31	.62
Pride	3.46	0.61	−0.28	.51	3.55	0.67	−0.29	.49	3.43	0.72	−0.36	.41
Anger	1.93	0.69	0.74	.59	1.74	0.72	1.11	.59	1.85	0.77	0.97	.57
Anxiety	2.31	0.69	0.28	.55	2.02	0.75	0.59	.51	2.38	0.88	0.47	.56
Shame	2.30	0.79	0.36	.62	2.22	0.93	0.58	.65	2.68	1.08	0.24	.68
Hopelessness	1.77	0.67	0.99	.66	1.67	0.72	1.10	.63	1.99	0.93	0.95	.67
Boredom	2.80	0.89	0.07	.71	2.90	1.00	0.07	.73	3.17	0.92	−0.14	.62
<i>Learning Settings</i>												
Enjoyment	3.31	0.58	−0.23	.45	3.52	0.67	−0.29	.43	3.51	0.73	−0.42	.48
Hope	3.38	0.61	−0.11	.52	3.31	0.67	0.03	.56	3.22	0.83	−0.14	.60
Pride	3.60	0.67	−0.38	.48	3.64	0.71	−0.43	.48	3.51	0.76	−0.24	.47
Anger	2.45	0.78	0.16	.59	2.62	0.90	0.02	.58	2.72	0.87	0.21	.48
Anxiety	2.80	0.71	−0.24	.52	2.77	0.83	−0.05	.48	3.23	0.93	−0.27	.52
Shame	2.46	0.76	0.31	.55	2.26	0.84	0.28	.53	2.58	0.93	0.20	.53
Hopelessness	2.09	0.74	0.58	.63	2.06	0.83	0.74	.61	2.46	1.00	0.37	.61
Boredom	2.79	0.84	−0.08	.67	3.00	1.00	−0.12	.70	3.12	0.98	−0.06	.67
<i>Test Settings</i>												
Enjoyment	2.83	0.60	0.02	.45	2.61	0.75	0.07	.43	2.48	0.81	0.32	.51
Hope	3.23	0.62	0.11	.52	3.16	0.74	0.01	.58	2.94	0.89	−0.08	.65
Pride	3.13	0.65	−0.15	.58	3.16	0.74	−0.14	.53	3.00	0.81	−0.10	.56
Relief	3.59	0.71	−0.40	.52	3.61	0.75	−0.38	.48	3.73	0.81	−0.58	.59
Anger	2.34	0.73	0.26	.57	2.06	0.81	0.69	.54	2.09	0.78	0.59	.47
Anxiety	3.02	0.83	−0.02	.61	3.18	0.94	−0.02	.57	3.51	0.92	−0.46	.56
Shame	2.19	0.75	0.45	.61	1.81	0.78	0.84	.66	2.19	0.89	0.63	.61
Hopelessness	2.01	0.76	0.64	.69	1.94	0.85	0.93	.69	2.40	1.02	0.47	.68

Note. $\bar{r}_{i(t-i)}$ denotes the average part-whole corrected item-total correlation. Items were answered on Likert scales (1 = *strongly disagree*, 5 = *strongly agree*). $N = 389$ for the AEQ and the AEQ-S in the reanalysis, $N = 471$ for the AEQ-S in the validation study.

Table 2

Comparison of the AEQ scales and the AEQ-S scales regarding their reliability and correlations between AEQ and AEQ-S scales in the reanalysis of Pekrun et al. (2011).

	AEQ				AEQ-S					
	Items	α	Items	$\Delta\alpha$	Reanalysis				Validation	
					α	$\Delta\alpha$	r	r^*	α	$\Delta\alpha$
<i>Class Setting</i>										
Enjoyment	10	.85	4	6	.75	.10	.93	.81	.73	.12
Hope	8	.79	4	4	.73	.06	.91	.75	.81	-.02
Pride	9	.81	4	5	.70	.11	.91	.76	.63	.18
Anger	9	.86	4	5	.78	.08	.91	.81	.75	.11
Anxiety	12	.86	4	8	.71	.15	.88	.78	.75	.11
Shame	11	.89	4	7	.83	.06	.92	.85	.84	.05
Hopelessness	10	.90	4	6	.81	.09	.94	.86	.84	.06
Boredom	11	.93	4	7	.88	.05	.95	.90	.80	.13
<i>Average</i>	10.0	.86	4.0	6.0	.77	.09	.92	.82	.77	.09
<i>Learning Setting</i>										
Enjoyment	10	.78	4	6	.64	.14	.89	.72	.69	.09
Hope	6	.77	4	2	.76	.01	.94	.76	.79	-.02
Pride	6	.75	4	2	.70	.05	.95	.74	.68	.07
Anger	9	.86	4	5	.77	.09	.92	.81	.70	.16
Anxiety	11	.84	4	7	.69	.15	.90	.77	.72	.12
Shame	11	.86	4	7	.73	.13	.91	.80	.73	.13
Hopelessness	11	.90	4	7	.79	.11	.93	.85	.80	.10
Boredom	11	.91	4	7	.85	.06	.95	.88	.83	.08
<i>Average</i>	9.4	.83	4.0	5.4	.74	.09	.93	.80	.74	.09
<i>Test Setting</i>										
Enjoyment	10	.78	4	6	.66	.12	.88	.71	.72	.06
Hope	8	.81	4	4	.77	.04	.92	.78	.82	-.01
Pride	10	.86	4	6	.74	.12	.93	.81	.76	.10
Relief	6	.77	4	2	.70	.07	.96	.74	.78	-.01
Anger	10	.86	4	6	.74	.12	.92	.80	.69	.17
Anxiety	12	.90	4	8	.77	.13	.93	.84	.76	.14
Shame	10	.87	4	6	.83	.04	.90	.82	.80	.07
Hopelessness	11	.92	4	7	.85	.07	.94	.88	.84	.08
<i>Average</i>	9.6	.85	4.0	5.6	.76	.09	.93	.81	.77	.08
<i>Overall Average</i>	9.7	.85	4.0	5.7	.76	.09	.92	.81	.76	.09

Note. α denotes Cronbach's alpha. $\Delta\alpha$ is the difference in α between AEQ and AEQ-S scales. r denotes the Pearson product-moment correlation. r^* is r adjusted for overlapping error variance according to Levy (1967). Average correlations were computed based on Fisher's z-transformed correlation coefficients. Items were answered on Likert scales (1 = *strongly disagree*, 5 = *strongly agree*). $N = 389$ for the AEQ and the AEQ-S in the reanalysis, $N = 471$ for the AEQ-S in the validation study.

Fisher's z-transformation). Because the AEQ-S scales in the study based on Pekrun et al. (2011) were embedded in the full AEQ scales, we also computed Levy-corrected correlations that account for shared error variance (Levy, 1967). These correlations ranged from $r^* = .75$ to .90 for class-related, $r^* = .72$ to .88 for learning-related, and $r^* = .71$ to .88 for test-related emotions. The average Levy-corrected correlation was $M_{r^*} = .81$.

3.3. Nomological network

3.3.1. Structural relations between achievement emotions

We estimated the four CFA models Pekrun et al. (2011) used to establish the relations between the achievement emotions assessed with the AEQ: a model with one general factor across emotions and settings (Model A), a model with nine factors representing emotions (Model B), a model with three factors representing settings (Model C), and a model with nine factors representing emotions plus correlated uniqueness within settings (Model D). We replicated Pekrun et al.'s findings with respect to these CFA models for the AEQ-S. Model D showed reasonable fit (based on recommendations of Hu & Bentler, 1999, for interpreting fit indices) both in the reanalysis, $\chi^2(133) = 340.82$, CFI = 0.96, TLI = 0.92, SRMR = 0.05, RMSEA = 0.063, and in the validation study, $\chi^2(133) = 427.56$, CFI = 0.96, TLI = 0.91, SRMR = 0.05, RMSEA = 0.069. Moreover, Model D provided substantially better fit than Models A/B/C, which performed consistently poorly in the reanalysis, $\chi^2(252/217/249) = 2389.95/1148.71/2231.97$, CFI = 0.58/0.82/0.61, TLI = 0.55/0.77/0.57, SRMR = 0.13/0.07/0.12, RMSEA = 0.148/0.105/0.143, and in the validation study, $\chi^2(252/217/249) = 3105.67/$

1308.85/3015.45, CFI = 0.59/0.84/0.60, TLI = 0.55/0.80/0.55, SRMR = 0.12/0.06/0.12, RMSEA = 0.155/0.103/0.154.

3.3.2. Correlations of emotions and settings

Table 3 shows the correlations of emotions within settings (e.g., the correlation between enjoyment and hope in class settings) for the AEQ scales (2011) with minor differences due to the above-mentioned differences in analyses. Table 4 shows the equivalent correlations for the AEQ-S scales based on the reanalysis (lower diagonal) and the validation study (upper diagonal). Corroborating the CFA analysis, the patterns of correlations obtained with the AEQ-S were in general highly similar to the pattern of correlations obtained with the AEQ, both in the reanalysis of the Pekrun et al. (2011) dataset and in the validation study. This is also captured by Fig. 2, which additionally reveals that some correlations between negative emotions tended to be lower in the AEQ-S than in the AEQ (e.g., correlations between anxiety and shame). The correlations of settings within emotions (e.g., correlation between enjoyment in class and enjoyment in learning settings) are presented in Table 5. Again, the pattern of correlations obtained with the AEQ-S was in general very similar to the pattern of correlations obtained with the AEQ, both in the reanalysis of the Pekrun et al. (2011) dataset and in the validation study. Fig. 3 shows this correspondence between the instruments, while again revealing that correlations involving negative emotions tended to be smaller in the AEQ-S than in the AEQ. There were only few notable exceptions from this general pattern; for instance, the association between class- and learning-related hope was stronger in the AEQ-S based on the validation study than in the AEQ.

Overall, structural relations and correlational patterns of emotions

Table 3

Correlations of emotions within settings of the full AEQ scales (Pekrun et al., 2011, dataset).

Emotion	Setting	1	2	3	4	5	6	7	8
1 Enjoyment									
2 Hope	Class	.71							
	Learning	.64							
	Test	.70							
3 Pride	Class	.62	.68						
	Learning	.72	.64						
	Test	.71	.68						
4 Relief	Class	—	—	—					
	Learning	—	—	—					
	Test	.06	-.03	.23					
5 Anger	Class	-.41	-.35	-.23	—				
	Learning	-.44	-.53	-.33	—				
	Test	-.26	-.37	-.21	.08				
6 Anxiety	Class	-.24	-.36	-.16	—	.64			
	Learning	-.12	-.42	-.15	—	.61			
	Test	-.39	-.49	-.30	.36	.58			
7 Shame	Class	-.27	-.33	-.20	—	.58	.79		
	Learning	-.15	-.43	-.23	—	.57	.69		
	Test	-.29	-.43	-.36	.06	.64	.66		
8 Hopelessness	Class	-.35	-.45	-.26	—	.77	.70	.63	
	Learning	-.32	-.58	-.42	—	.67	.68	.75	
	Test	-.39	-.52	-.39	.00	.72	.66	.77	
9 Boredom	Class	-.57	-.42	-.28	—	.62	.46	.41	.50
	Learning	-.51	-.48	-.39	—	.76	.50	.50	.58
	Test	—	—	—	—	—	—	—	—

Note. The AEQ assesses relief only in test-related settings and boredom only in class-related and learning-related settings. $N = 389$.

Table 4

Correlations of emotions within settings of the AEQ-S scales in the reanalysis of Pekrun et al., 2011 (lower diagonal) and in the validation study (upper diagonal).

Emotion	Setting	1	2	3	4	5	6	7	8	9
1 Enjoyment	Class		.67	.56	—	-.31	-.30	-.25	-.34	-.54
	Learning		.64	.63	—	-.33	-.08	-.20	-.30	-.43
	Test		.53	.60	.06	.00	-.32	-.25	-.34	—
2 Hope	Class	.64		.73	—	-.37	-.43	-.40	-.52	-.42
	Learning	.53		.68	—	-.39	-.42	-.46	-.60	-.42
	test	.54		.69	.18	-.11	-.51	-.48	-.64	—
3 Pride	Class	.53	.58		—	-.26	-.28	-.27	-.40	-.27
	Learning	.69	.59		—	-.31	-.20	-.33	-.44	-.38
	Test	.42	.55		.30	-.02	-.35	-.39	-.51	—
4 Relief	Class	—	—	—	—	—	—	—	—	—
	Learning	—	—	—	—	—	—	—	—	—
	Test	-.09	-.10	.17		-.04	.16	-.15	-.13	—
5 Anger	Class	-.40	-.33	-.26	—		.50	.40	.64	.42
	Learning	-.41	-.41	-.34	—		.46	.41	.52	.63
	Test	-.17	-.26	-.17	.00		.26	.39	.38	—
6 Anxiety	Class	-.28	-.39	-.23	—	.64		.63	.67	.30
	Learning	-.17	-.41	-.19	—	.51		.56	.64	.33
	Test	-.37	-.49	-.19	.38	.39		.51	.58	—
7 Shame	Class	-.32	-.32	-.24	—	.55	.64		.52	.29
	Learning	-.23	-.44	-.27	—	.49	.59		.74	.33
	Test	-.10	-.35	-.31	-.02	.60	.47		.68	—
8 Hopelessness	Class	-.34	-.38	-.27	—	.71	.64	.58		.34
	Learning	-.35	-.52	-.42	—	.50	.58	.73		.40
	Test	-.29	-.50	-.37	-.03	.62	.54	.76		—
9 Boredom	Class	-.52	-.31	-.28	—	.48	.37	.34	.40	
	Learning	-.49	-.37	-.40	—	.71	.44	.44	.46	
	Test	—	—	—	—	—	—	—	—	—

Note. The AEQ-S assesses relief only in test-related settings and boredom only in class-related and learning-related settings. $N = 389$ for the reanalysis, $N = 471$ for the validation study.

and settings observed in the AEQ and the AEQ-S corresponded well, including cross-validation in a new and independent study, indicating that the AEQ-S scales represent a similar nomological network as the full AEQ scales. In particular, the nature of the associations (i.e., positive or negative) were fully retained, and their magnitudes were predominantly similar. Only scales involving negative emotions tended to produce weaker associations in the AEQ-S than in the AEQ, maybe reflecting that these scales are quite long in the AEQ and were thus shortened more than scales measuring positive emotions.

3.3.3. Correlations with external measures

Table 6 shows the correlations of the AEQ scales and external measures (i.e., antecedents and consequences of academic emotions) reported by Pekrun et al. (2011). Tables 7 and 8 show the corresponding correlations between the short AEQ scales and the external measures based on the Pekrun et al. (2011) data and the validation data, respectively. Again, we observed substantial similarities between the patterns of correlations of the AEQ and the AEQ-S with external measures (for a summary display, see Fig. 4). There were a few deviations from this rule,

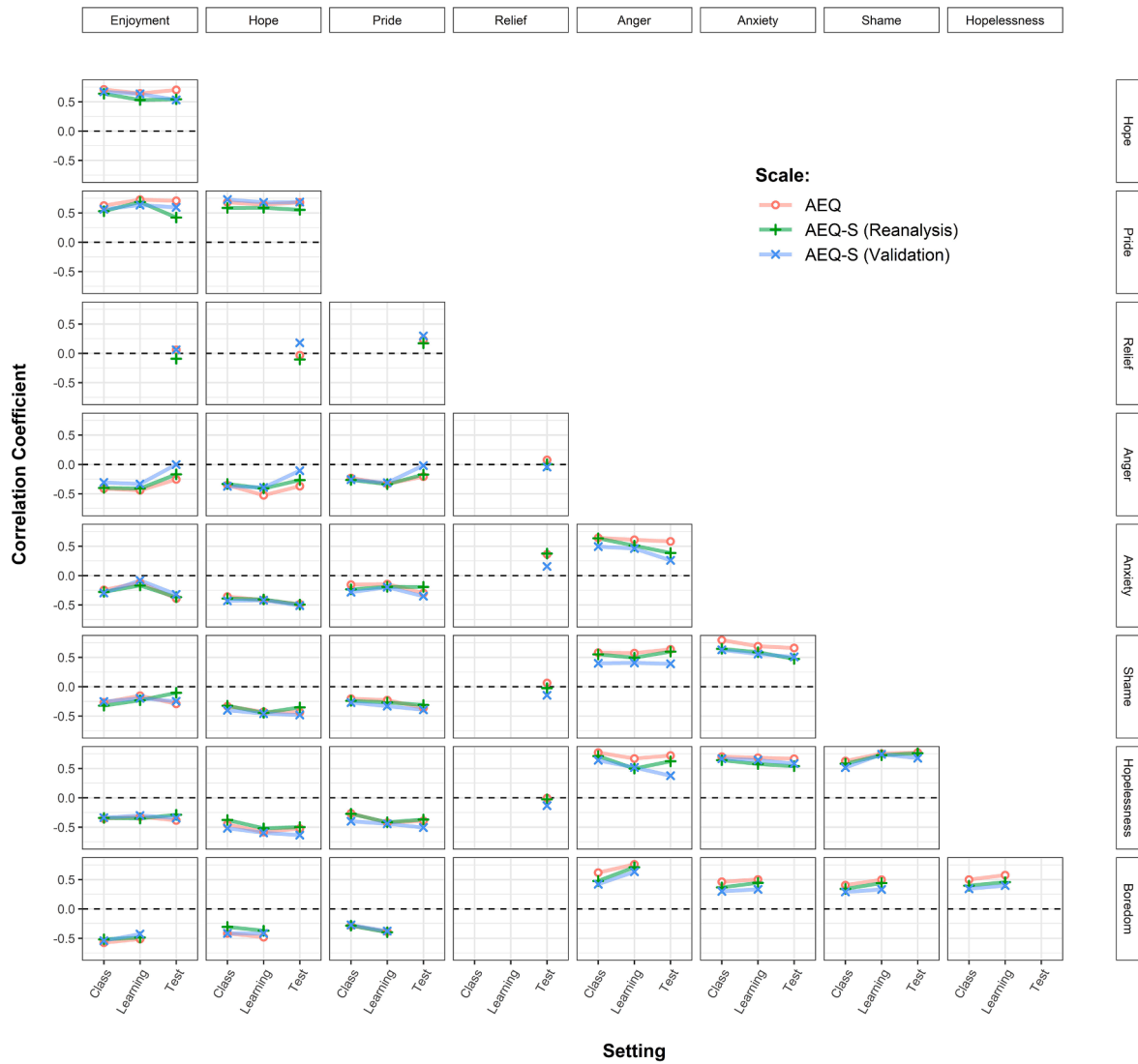


Fig. 2. Correlational patterns of emotions within each setting covered by the AEQ and the AEQ-S (based on the reanalysis and the validation study).

which were usually limited to the validation study (e.g., the associations between anger in test settings, on the one hand, and perceived academic control and self-efficacy, on the other). Overall, the results suggest that the nomological network of achievement emotions and their antecedents and consequences are highly similar for the AEQ and the AEQ-S.

4. Discussion

The Achievement Emotions Questionnaire (AEQ) has pioneered research on achievement emotions in academic settings and is a popular and well-established instrument in educational research and beyond. Its most distinctive feature is the comprehensive measurement of various achievement emotions with their affective, cognitive, motivational, and physiological components across academic settings. However, this broad scope is achieved with relatively large numbers of items that can make it difficult or even impossible to administer the full AEQ scales in empirical studies. So far, researchers had to solve this problem in an *ad hoc* manner by shortening the scales to fit their purpose. This is not a satisfactory state of affairs because shortening scales *ad hoc* is not sufficiently based on an inspection of psychometric quality and may fail to render generalizable results. In the present research, we therefore systematically developed the AEQ-S as a short form of the AEQ and investigated its psychometric properties.

4.1. Reliability

We observed satisfactory reliabilities of the AEQ-S across class-, learning-, and test-related emotions, albeit these reliabilities were lower than reliabilities of the AEQ. This is, however, not surprising given that shortening questionnaires almost inevitably reduces their reliability (Nunnally & Bernstein, 1994). It should also be noted that reliability is not the only criterion one should consider when developing short scales. While selecting those items from the full scales that have the highest item-total correlations would maximize the reliability of the resulting short scales, at the same time this could inadvertently narrow the content domain in comparison to the full scales (Smith et al., 2000). In case of the AEQ, we deemed it important to retain the conceptual scope as far as possible, for instance, by making sure that the AEQ-S retained items for each emotional component (i.e., affective, cognitive, motivational, and physiological) covered by the AEQ. Moreover, the average reduction in scale length of about 40% likely justifies the 10% drop in reliability in many applications. Specifically, researchers who want to measure all achievement emotions in all academic settings covered by the AEQ can achieve this in 20 min with the AEQ-S, compared to the approximately 50 min that are required for administering the full AEQ scales (Pekrun, Goetz, & Perry, 2005).

That said, we want to emphasize that we do not consider the AEQ-S

Table 5

Correlations of settings within emotions for the AEQ and the AEQ-S scales.

	AEQ	AEQ-S	
		Reanalysis	Validation
<i>Class vs. learning</i>			
Enjoyment	.61	.50	.57
Hope	.52	.48	.76
Pride	.59	.62	.72
Relief	—	—	—
Anger	.61	.48	.50
Anxiety	.66	.47	.58
Shame	.71	.60	.58
Hopelessness	.75	.65	.72
Boredom	.73	.66	.61
<i>Class vs. test</i>			
Enjoyment	.47	.34	.41
Hope	.58	.49	.63
Pride	.60	.56	.54
Relief	—	—	—
Anger	.73	.69	.55
Anxiety	.63	.41	.49
Shame	.70	.59	.52
Hopelessness	.79	.72	.68
Boredom	—	—	—
<i>Learning vs. test</i>			
Enjoyment	.58	.43	.52
Hope	.64	.64	.72
Pride	.60	.57	.56
Relief	—	—	—
Anger	.68	.54	.45
Anxiety	.74	.59	.62
Shame	.78	.68	.66
Hopelessness	.82	.73	.73
Boredom	—	—	—

Note. $N = 389$ for the AEQ and the AEQ-S in the reanalysis, $N = 471$ for the AEQ-S in the validation study.

as a replacement of the AEQ. Rather, the two instruments provide different solutions for how to balance the inherent tradeoff between reliability and parsimony in the measurement of achievement emotions. As a consequence, researchers will likely continue to use the AEQ when the reliable measurement of achievement emotions lies at the core of their investigation, while they might resort to the AEQ-S when brevity is crucial and outweighs the loss in reliability.

4.2. Information reproduction

Besides being reliable, short scales should also reproduce the information obtained by the full scales (Gogol et al., 2014). We checked this in the reanalysis of the Pekrun et al. (2011) data by examining correlations between the AEQ scales and the AEQ-S scales. These correlations turned out to be sizable, and all of them remained larger than 0.70 even after we had accounted for shared error variance (Girard & Christensen, 2008; Levy, 1967). As such, the findings suggest that the AEQ-S can reproduce the information that is obtained with the AEQ.

4.3. Nomological network

A final criterion for short scales is that they are embedded in the same nomological network as the full scales (Gogol et al., 2014). We tested this requirement in several ways both in the reanalysis of the Pekrun et al. (2011) data and with the validation data. Confirmatory factor analyses showed that a two-facet model (emotions $\times$ setting facets) described the data from the AEQ-S best, closely mirroring the structural relations between achievement emotions that has been established for the AEQ (Pekrun et al., 2011). The patterns of correlations of emotion scales within settings (e.g., class-related enjoyment vs. hope), of settings within emotion scales (e.g., class-related vs. learning-related enjoyment), and of emotion scales and external measures of antecedents and consequences of achievement emotions (e.g., class-related enjoyment and perceived control) were highly similar for the AEQ and the AEQ-S scales. Taken together, our results suggest that the AEQ and the AEQ-S scales are embedded in a similar nomological network.

4.4. Further considerations

For our validation study, we drew upon a sample of English-speaking Canadian university students similar to the sample recruited by Pekrun et al. (2011). This was done to maximize comparability of the data underlying the AEQ and the AEQ-S in this study and to avoid differences in sample characteristics that could jeopardize the comparability of the results. On the downside, the present validation leaves it open to question whether the findings can be generalized beyond samples with these characteristics. For the AEQ-Mathematics, cross-cultural measurement equivalence has been demonstrated (Frenzel, Thrash, Pekrun, & Goetz, 2007). Future research should examine reliability, validity, and

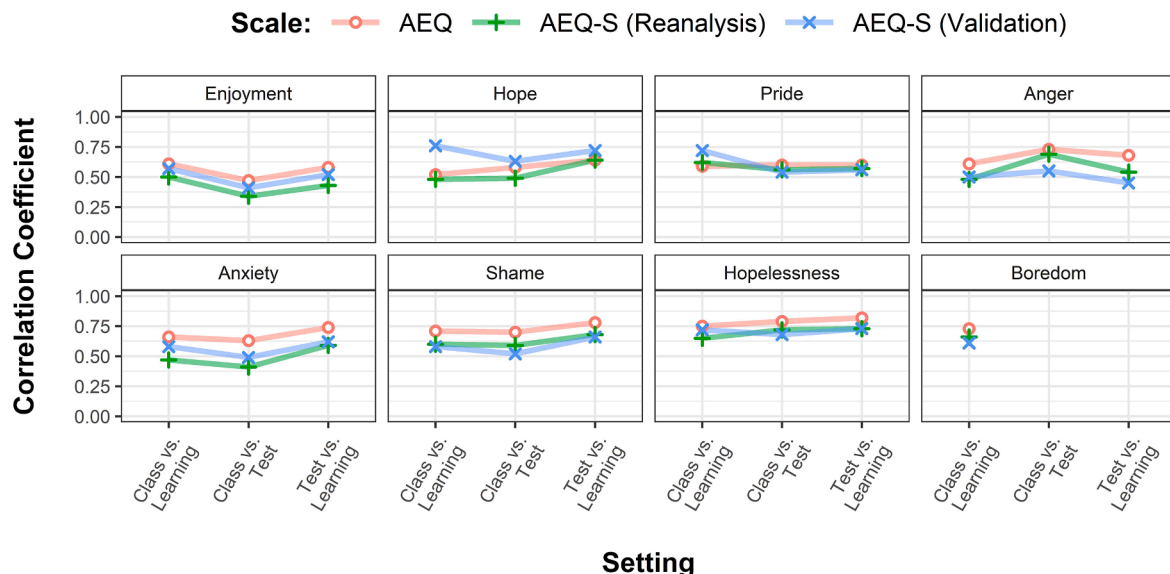


Fig. 3. Correlational patterns of settings within each emotion measured by the AEQ and the AEQ-S (based on the reanalysis and the validation study).

Table 6
Correlations between full AEQ scales (Pekrun et al., 2011) and external measures.

Emotion	Setting	Antecedents			Consequences							Academic performance	Gender
		Academic control	Self-efficacy	Task value	Intrinsic goal orientation	Extrinsic goal orientation	Effort regulation	Elaboration	Rehearsal	Self-regulation	External regulation		
Enjoyment	Class	.32	.39	.56	.44	.15	.29	.41	.19	.27	−.07	.18	−.14
	Learning	.30	.45	.51	.49	.22	.37	.41	.18	.34	−.01	.19	−.07
	Test	.18	.51	.35	.37	.12	.37	.38	.23	.38	−.02	.27	.08
Hope	Class	.39	.55	.46	.41	.17	.38	.44	.24	.45	−.03	.22	−.04
	Learning	.42	.56	.41	.43	.07	.47	.40	.09	.52	−.14	.33	.09
	Test	.32	.61	.36	.40	.05	.45	.41	.18	.52	−.05	.29	.12
Pride	Class	.35	.51	.45	.34	.34	.36	.42	.32	.43	.10	.16	−.07
	Learning	.42	.49	.44	.40	.26	.42	.42	.29	.46	.03	.26	−.05
	Test	.29	.56	.32	.34	.17	.42	.44	.31	.49	.06	.33	.06
Relief	Test	.17	.06	.15	.05	.25	.04	.21	.19	.14	.22	.16	−.07
Anger	Class	−.60	−.34	−.45	−.22	−.06	−.38	−.31	−.11	−.26	.27	−.28	.13
	Learning	−.39	−.39	−.33	−.27	.02	−.43	−.29	−.01	−.31	.30	−.24	.01
	Test	−.52	−.39	−.31	−.20	.08	−.37	−.29	−.03	−.30	.34	−.30	.03
Anxiety	Class	−.46	−.38	−.17	−.10	.17	−.33	−.17	.02	−.29	.32	−.20	−.07
	Learning	−.28	−.34	−.08	−.11	.20	−.30	−.13	.08	−.30	.35	−.13	−.15
	Test	−.30	−.38	−.14	−.16	.26	−.28	−.12	.11	−.28	.32	−.15	−.17
Shame	Class	−.45	−.34	−.21	−.09	.14	−.31	−.18	.02	−.26	.23	−.17	−.07
	Learning	−.38	−.35	−.15	−.07	.18	−.41	−.19	.03	−.34	.30	−.27	−.08
	Test	−.42	−.42	−.18	−.08	.19	−.37	−.25	−.02	−.37	.22	−.37	−.09
Hopelessness	Class	−.66	−.47	−.41	−.25	−.03	−.43	−.41	−.13	−.37	.25	−.35	.08
	Learning	−.59	−.51	−.29	−.26	.10	−.45	−.35	−.03	−.45	.32	−.31	−.04
	Test	−.58	−.50	−.34	−.22	.08	−.44	−.35	−.04	−.40	.25	−.34	−.03
Boredom	Class	−.29	−.28	−.38	−.23	.00	−.42	−.19	−.04	−.16	.24	−.16	.06
	Learning	−.31	−.34	−.38	−.26	−.02	−.48	−.25	−.04	−.28	.24	−.22	.08

Note. $N = 389$.

Table 7
Correlations between AEQ-S scales and external measures in the reanalysis of [Pekrun et al. \(2011\)](#).

Emotion	Setting	Antecedents			Consequences							Academic performance	Gender
		Academic control	Self-efficacy	Task value	Intrinsic goal orientation	Extrinsic goal orientation	Effort regulation	Elaboration	Rehearsal	Self-regulation	External regulation		
Enjoyment	Class	.31	.34	.52	.40	.11	.27	.35	.16	.25	−.06	.16	−.14
	Learning	.33	.45	.48	.44	.18	.40	.39	.18	.37	−.04	.24	−.09
	Test	.10	.40	.30	.33	.01	.25	.27	.12	.24	−.11	.17	.11
Hope	Class	.37	.48	.42	.37	.15	.29	.38	.19	.38	−.01	.18	−.05
	Learning	.42	.57	.34	.39	.02	.44	.39	.08	.53	−.15	.34	.12
	Test	.30	.60	.30	.34	−.04	.37	.31	.07	.45	−.04	.26	.19
Pride	Class	.38	.49	.47	.36	.31	.37	.40	.31	.39	.05	.10	−.07
	Learning	.44	.53	.46	.39	.28	.44	.45	.30	.46	.02	.25	−.05
	Test	.30	.51	.33	.31	.19	.38	.43	.31	.45	.06	.31	.01
Relief	Test	.16	.07	.13	.04	.23	.02	.18	.17	.12	.23	.18	−.07
Anger	Class	−.60	−.30	−.41	−.20	−.09	−.33	−.30	−.13	−.23	.20	−.25	.13
	Learning	−.32	−.31	−.30	−.23	.03	−.38	−.25	−.02	−.25	.24	−.18	.03
	Test	−.53	−.36	−.34	−.21	.03	−.34	−.32	−.06	−.28	.27	−.26	.06
Anxiety	Class	−.50	−.34	−.26	−.12	.06	−.32	−.20	−.02	−.25	.22	−.20	−.01
	Learning	−.25	−.33	−.10	−.12	.19	−.32	−.12	.05	−.30	.31	−.08	−.11
	Test	−.23	−.34	−.10	−.13	.25	−.20	−.09	.11	−.25	.30	−.11	−.15
Shame	Class	−.45	−.32	−.21	−.14	.09	−.28	−.22	.00	−.24	.19	−.16	−.03
	Learning	−.44	−.38	−.23	−.12	.10	−.38	−.24	.01	−.36	.30	−.31	−.04
	Test	−.48	−.39	−.23	−.08	.07	−.33	−.31	−.06	−.38	.16	−.34	−.01
Hopelessness	Class	−.61	−.44	−.42	−.24	−.05	−.42	−.37	−.16	−.35	.19	−.31	.08
	Learning	−.58	−.48	−.25	−.22	.09	−.41	−.37	−.08	−.44	.30	−.28	−.03
	Test	−.55	−.47	−.32	−.23	.03	−.40	−.37	−.06	−.40	.21	−.32	−.02
Boredom	Class	−.25	−.22	−.35	−.22	.00	−.38	−.16	−.05	−.11	.25	−.10	.08
	Learning	−.26	−.31	−.34	−.26	−.05	−.46	−.24	−.05	−.25	.26	−.17	.07

Note. $N = 389$.

Table 8
Correlations between AEQ-S scales and external measures in the validation study.

Emotion	Setting	Antecedents			Consequences							Academic performance	Gender
		Academic control	Self-efficacy	Task value	Intrinsic goal orientation	Extrinsic goal orientation	Effort regulation	Elaboration	Rehearsal	Self-regulation	External regulation		
Enjoyment	Class	.14	.39	.59	.46	.07	.30	.37	.21	.36	−.01	.22	.04
	Learning	.27	.46	.58	.50	.23	.38	.40	.28	.52	.07	.20	−.03
	Test	.22	.51	.29	.36	.04	.30	.21	.11	.40	−.05	.26	.23
Hope	Class	.31	.67	.48	.41	.08	.41	.34	.14	.53	−.01	.28	.10
	Learning	.33	.72	.44	.41	.06	.43	.30	.16	.57	.02	.26	.16
	Test	.31	.64	.25	.30	−.10	.32	.15	.05	.41	−.01	.19	.32
Pride	Class	.27	.53	.50	.37	.22	.37	.38	.20	.50	.01	.21	.03
	Learning	.33	.51	.38	.29	.22	.50	.34	.25	.54	.00	.26	−.02
	Test	.30	.57	.26	.22	.02	.33	.21	.10	.41	.00	.28	.26
Relief	Test	.18	.16	.10	.04	.14	.10	.12	.13	.15	.14	.08	.06
Anger	Class	−.37	−.23	−.31	−.18	.01	−.38	−.15	−.07	−.16	.09	−.22	.08
	Learning	−.26	−.32	−.33	−.22	.02	−.45	−.16	.00	−.29	.12	−.26	.07
	Test	−.31	−.13	−.24	−.17	.10	−.24	−.10	.02	−.12	.10	−.16	.15
Anxiety	Class	−.36	−.39	−.22	−.19	.13	−.32	−.11	.14	−.23	.12	−.25	−.11
	Learning	−.23	−.40	−.04	−.09	.21	−.28	.05	.17	−.22	.12	−.27	−.27
	Test	−.21	−.38	−.03	−.12	.31	−.15	.04	.20	−.13	.16	−.17	−.21
Shame	Class	−.22	−.34	−.12	−.13	.15	−.24	−.04	.06	−.23	.09	−.18	−.18
	Learning	−.43	−.55	−.22	−.23	.15	−.43	−.14	.07	−.35	.16	−.34	−.11
	Test	−.43	−.43	−.12	−.18	.17	−.33	−.09	.02	−.30	.11	−.32	−.09
Hopelessness	Class	−.52	−.50	−.33	−.24	.00	−.52	−.18	.00	−.34	.15	−.35	−.03
	Learning	−.49	−.61	−.29	−.25	.09	−.49	−.14	.07	−.41	.17	−.37	−.14
	Test	−.41	−.54	−.15	−.20	.11	−.40	−.09	.02	−.34	.11	−.33	−.17
Boredom	Class	−.12	−.22	−.36	−.27	.05	−.30	−.17	.00	−.22	.09	−.14	.00
	Learning	−.15	−.29	−.33	−.28	.01	−.50	−.27	−.05	−.39	.16	−.27	.05

Note. $N = 471$.

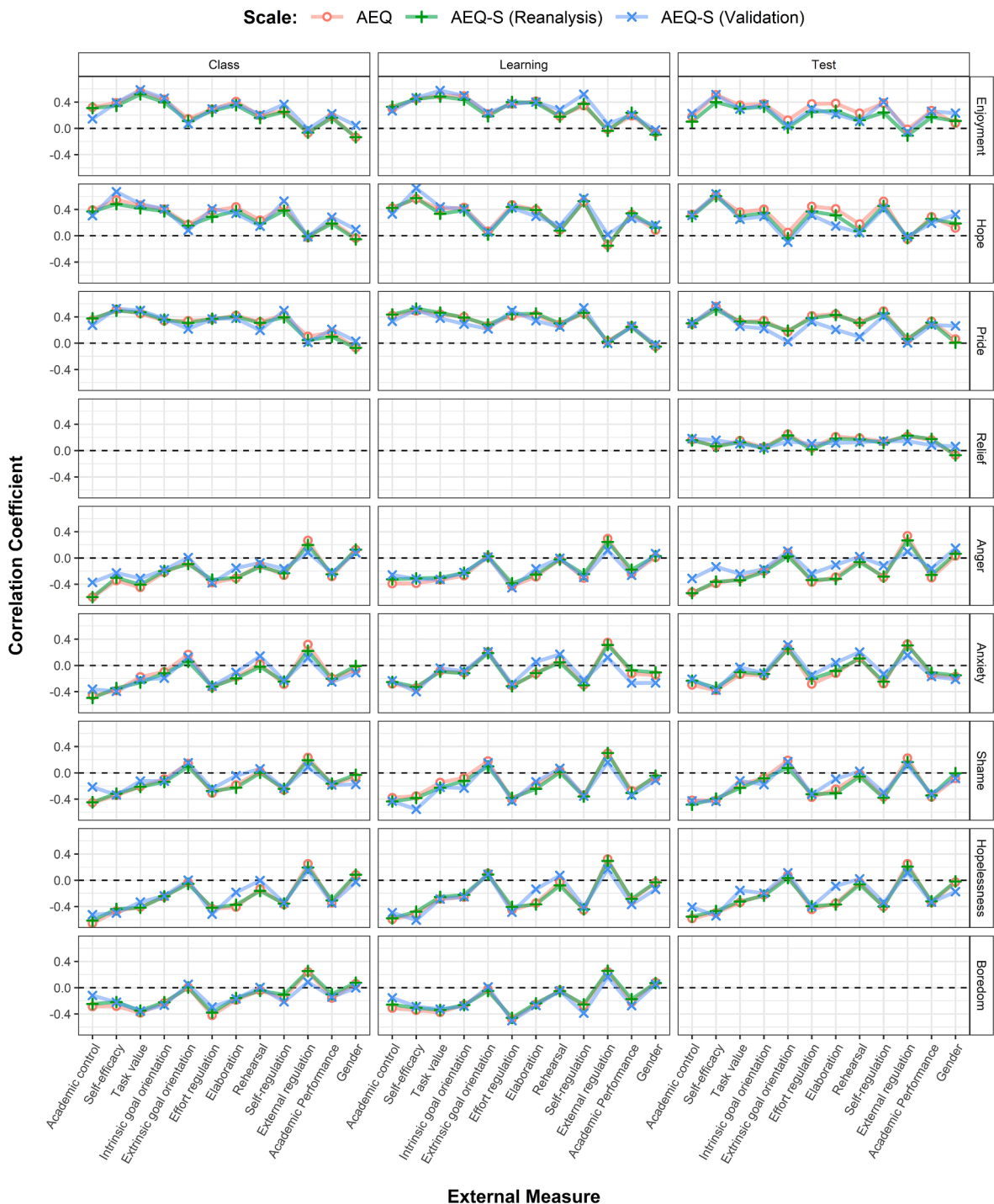


Fig. 4. Correlational patterns of emotions and external measures of their antecedents and consequences for the AEQ and the AEQ-S (based on the reanalysis and the validation study).

measurement equivalence of the AEQ-S across different populations and languages as well. Moreover, to facilitate comparisons between the AEQ-S and the AEQ we measured the antecedents and consequences of emotions in the validation study in the same way as in the study by Pekrun et al. (2011). It would be desirable to further validate the AEQ-S using alternative measures, for instance, by replacing self-reported GPA with more objective assessments of academic performance or by using more reliable measures of motivation and self-regulation than those provided by the MSLQ. Beyond showing agreement between the AEQ and the AEQ-S, this would further contribute to establishing meaningful associations between the AEQ-S and external measures.

The assessment of information reproduction of the AEQ-S in the reanalysis of the Pekrun et al. (2011) data rests on their correlations with the AEQ using participants who worked on the full AEQ scales. This raises the issue of overlapping error variance, which we accounted for using a correction formula suggested by Levy (1967). Still, it would be even better to avoid the issue altogether by giving participants both the AEQ and the AEQ-S and then correlate their answers to both instruments (Smith et al., 2000). To implement such a procedure, for instance, Pfeiffer, Hagemann, and Backenstrass (2011) suggest to assign participants to different groups that receive specific combinations of the short and full scales of a questionnaire, which allows to experimentally

control for variance overlap.

Our analysis of studies using the AEQ revealed not only an interest in shorter scales but also in single-item AEQ scales which are, for instance, of crucial importance in experience sampling designs (Goetz, Bieg, & Hall, 2016). Indeed, it has been demonstrated that single-item scales can be reliable and valid instruments for assessing motivational and emotional constructs (Gogol et al., 2014; Wanous, Reichers, & Hudy, 1997). Therefore, another route for future research is to systematically develop a single-item version of the AEQ and assess its psychometric characteristics. However, single-item scales would likely benefit from relying on items that directly ask for the emotion at hand (e.g., “How much enjoyment do you experience?”) rather than using the original AEQ items, which are tailored to specific emotion components (e.g., the affective component).

5. Conclusion

In the present research, we introduced the AEQ-S as a short form of

the AEQ. The need for short scales is evident in work that draws upon the AEQ (and is observable more generally as well; Ziegler, Kemper, & Kruyen, 2014). Our results show that the AEQ-S constitutes a reliable instrument that is able to reproduce the information obtained by the AEQ and represents the same nomological network. Thus, the AEQ-S can be used whenever administration time or space are limited, thereby continuing to advance the study of achievement emotions in a wide range of authentic educational contexts.

Acknowledgements

We gratefully acknowledge the support of Kimberly Noels (University of Alberta, Canada) in recruiting participants for the research reported in this paper. We also thank Takuya Yanagida (University of Vienna, Austria) for helpful comments on the data analysis.

Appendix A. Instructions and items of the AEQ-S scales

Participants in the validation study responded to all items by using rating scales with five categories: 1 (*strongly disagree*) to 5 (*strongly agree*).

Class-related emotions

Attending classes at university can induce different feelings. The following questions refer to emotions you may experience when being in class at university. Before answering the questions, please recall some typical situations of being in class which you have experienced during the course of your studies. Please indicate how you feel, typically, when being in class. Please read each statement carefully and respond using the scale provided.

Enjoyment	I enjoy being in class. I am looking forward to learning a lot in this class. I am motivated to go to this class because it's exciting. I enjoy participating so much that I get energized.
Hope	I am confident when I go to class. I am full of hope. I am confident because I understand the material. Being confident that I will understand the material motivates me.
Pride	I am proud of myself. I think that I can be proud of what I know about this subject. Because I take pride in my accomplishments in this course, I am motivated to continue. When I do well in class, my heart throbs with pride.
Anger	I am angry. When I think of the time I waste in class I get aggravated. I wish I didn't have to attend class because it makes me angry. I feel anger welling up in me.
Anxiety	I feel nervous in class. Even before class, I worry whether I will be able to understand the material. Because I'm so nervous I would rather skip the class. I get tense in class.
Shame	I get embarrassed. When I say anything in class I feel like I am making a fool of myself. After I have said something in class I wish I could crawl into a hole and hide. Because I get embarrassed, I become tense and inhibited.
Hopelessness	I feel hopeless. I have lost all hope in understanding this class. Because I've given up, I don't have energy to go to class. I feel so hopeless all my energy is depleted.
Boredom	I get bored. The lecture bores me. I think about what else I might be doing rather than sitting in this boring class. I get restless because I can't wait for the class to end.

Learning-related emotions

Studying for your courses at university can induce different feelings. The following questions refer to emotions you may experience when studying. Before answering the questions, please recall some typical situations of studying which you have experienced during the course of your studies. Please indicate how you feel, typically, when studying. Please read each statement carefully and respond using the scale provided.

Enjoyment	I enjoy the challenge of learning the material. I enjoy dealing with the course material. I am so happy about the progress I made that I am motivated to continue studying. When my studies are going well, it gives me a rush.
Hope	I feel confident when studying. I feel confident that I will be able to master the material. I feel optimistic that I will make good progress at studying. My sense of confidence motivates me.
Pride	I'm proud of myself. I think I can be proud of my accomplishments at studying. Because I want to be proud of my accomplishments, I am very motivated. When I excel at my work, I swell with pride.
Anger	Studying makes me irritated. I get annoyed about having to study. I get so angry I feel like throwing the textbook out of the window. When I sit at my desk for a long time, my irritation makes me restless.
Anxiety	I get tense and nervous while studying. I worry whether I'm able to cope with all my work. While studying I feel like distracting myself in order to reduce my anxiety. Worry about not completing the material makes me sweat.
Shame	I feel ashamed. I feel ashamed when I realize that I lack ability. Because I have had so much trouble with the course material, I avoid discussing it. (*) When somebody notices how little I understand I avoid eye contact.
Hopelessness	I feel helpless. I'm resigned to the fact that I don't have the capacity to master this material. I feel so helpless that I can't give my studies my full efforts. My lack of confidence makes me exhausted before I even start.
Boredom	Studying for my courses bores me. The material is so boring that I find myself daydreaming. I would rather put off this boring work till tomorrow. While studying I seem to drift off because it's so boring.

(*) The original wording "much troubles" was changed to "much trouble" for language clarity.

Test-related emotions

Test and exams can induce different feelings. The following questions refer to emotions you may experience when taking tests or exams at university. Before answering the questions, please recall some typical situations of test-taking or exams which you have experienced during the course of your studies. Please indicate how you feel, typically, when taking a test or an exam. Please read each statement carefully and respond using the scale provided.

Enjoyment	I enjoy taking the exam. For me the test is a challenge that is enjoyable. Because I enjoy preparing for the test, I'm motivated to do more than is necessary. Before taking the exam, I sense a feeling of eagerness.
Hope	I am optimistic that everything will work out fine. I am very confident. I think about my exam optimistically. My confidence motivates me to prepare well.
Pride	I am proud of myself. I'm proud of how well I mastered the exam. Pride in my knowledge fuels my efforts in doing the test. After the exam I feel ten feet taller because I'm so proud.
Relief	After the exam I feel relief. (*) After the exam I feel freed. (*) After the exam the tension in my stomach is dissipated. (*) After the exam I finally can breathe easy again. (*)
Anger	I get angry. I get angry about the teacher's grading standards. I wish I could tell the teacher off.

(continued on next page)

(continued)

Anxiety	My anger makes the blood rush to my head.
	I am very nervous.
	I worry whether the test will be too difficult.
	I get so nervous I wish I could just skip the exam.
Shame	At the beginning of the test, my heart starts pounding.
	I feel ashamed.
	I get embarrassed because I can't answer the questions correctly.
	I get so embarrassed I want to run and hide.
Hopelessness	Because I am ashamed my pulse races.
	I feel hopeless.
	I start to think that no matter how hard I try I won't succeed on the test.
	I feel like giving up.
	I feel so resigned that I have no energy.

(*) A temporal specification ("After the exam") was added to these items.

References

- Frenzel, A. C., Thrash, T. M., Pekrun, R., & Goetz, T. (2007). Achievement emotions in Germany and China: A cross-cultural validation of the Academic Emotions Questionnaire – Mathematics. *Journal of Cross-Cultural Psychology*, 38, 302–309. <https://doi.org/10.1177/0022022107300276>.
- Girard, T. A., & Christensen, B. K. (2008). Clarifying problems and offering solutions for correlated error when assessing the validity of selected-subtest short forms. *Psychological Assessment*, 20, 76–80. <https://doi.org/10.1037/1040-3590.20.1.76>.
- Goetz, T. (2004). *Emotionales Erleben und selbstreguliertes Lernen bei Schülern im Fach Mathematik [Students' emotions and self-regulated learning in mathematics]*. München: Utz.
- Goetz, T., Bieg, M., & Hall, N. C. (2016). Assessing academic emotions via the experience sampling method. In M. Zembylas, & P. Schutz (Eds.), *Methodological advances in research on emotion in education* (pp. 279–298). New York: Springer.
- Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., & Preckel, F. (2014). "My questionnaire is too long!" The assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology*, 39, 188–205. <https://doi.org/10.1016/j.cedpsych.2014.04.002>.
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <https://doi.org/10.1080/10705519909540118>.
- King, R. B. (2010). What do students feel in school and how do we measure them? Examining the psychometric properties of the S-AEQ-F. *Philippine Journal of Psychology*, 43, 161–176.
- Levy, P. (1967). The correction for spurious correlation in the evaluation of short-form tests. *Journal of Clinical Psychology*, 23, 84–86. [https://doi.org/10.1002/1097-4679\(196701\)23:1<84::AID-JCLP2270230123>3.0.CO;2-2](https://doi.org/10.1002/1097-4679(196701)23:1<84::AID-JCLP2270230123>3.0.CO;2-2).
- Lichtenfeld, S., Pekrun, R., Stupnisky, R. H., Reiss, K., & Murayama, K. (2012). Measuring students' emotions in the early years: The Achievement Emotions Questionnaire-Elementary School (AEQ-ES). *Learning and Individual Differences*, 22, 190–201. <https://doi.org/10.1016/j.lindif.2011.04.009>.
- Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment*, 17, 81–102. <https://doi.org/10.1037/1040-3590.17.1.81>.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Peixoto, F., Mata, L., Monteiro, V., Sanches, C., & Pekrun, R. (2015). The achievement emotions questionnaire: Validation for pre-adolescent students. *European Journal of Developmental Psychology*, 12, 472–481. <https://doi.org/10.1080/17405629.2015.1040757>.
- Pekrun, R., Muis, K. R., Frenzel, A. C., & Goetz, T. (2017). *Emotions at school*. New York: Routledge.
- Pekrun, R., Goetz, T., & Frenzel, A. C. (2005). *Achievement Emotions Questionnaire – Mathematics (AEQ-M): User's manual*. Munich, Germany: University of Munich, Department of Psychology.
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36, 36–48. <https://doi.org/10.1016/j.cedpsych.2010.10.002>.
- Pekrun, R., Goetz, T., & Perry, R. P. (2005). *Achievement Emotions Questionnaire (AEQ) – User's manual*. Munich, Germany: Department of Psychology, University of Munich.
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37, 91–105. https://doi.org/10.1207/S15326985EP3702_4.
- Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, 88, 1653–1670. <https://doi.org/10.1111/cdev.12704>.
- Pekrun, R., Murayama, K., Marsh, H. W., Goetz, T., & Frenzel, A. C. (2019). Happy fish in little ponds: Testing a reference group model of achievement and emotion. *Journal of Personality and Social Psychology*, 117, 166–185. <https://doi.org/10.1037/pspp0000230>.
- Perry, R. P., Hladkyj, S., Pekrun, R., & Pelletier, S. (2001). Academic control and action control in the achievement of college students: A longitudinal field study. *Journal of Educational Psychology*, 93, 776–789. <https://doi.org/10.1037/0022-0663.93.4.776>.
- Pfeiffer, N., Hagemann, D., & Backenstrass, M. (2011). A new method for estimating the variance overlap between the short and the long form of a psychological test. *Educational and Psychological Measurement*, 71, 380–388. <https://doi.org/10.1177/0013164410386782>.
- Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. Ann Arbor, MI: University of Michigan.
- Raccanello, D. (2015). Students' expectations about interviewees' and interviewers' achievement emotions in job selection interviews. *Journal of Employment Counseling*, 52, 50–64. <https://doi.org/10.1002/joec.12004>.
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12, 102–111. <https://doi.org/10.1037/1040-3590.12.3.102>.
- Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology*, 82, 247–252.
- Zeidner, M. (1998). *Test anxiety – The state of the art*. New York, NY, USA: Plenum Press.
- Ziegler, M., Kemper, C. J., & Kruey, P. (2014). Short scales – Five misunderstandings and ways to overcome them. *Journal of Individual Differences*, 35, 185–189. <https://doi.org/10.1027/1614-0001/a000148>.