How Much Trait Variance Is Captured by Measures of Academic State Emotions?

A Latent State-Trait Analysis

Ulrike E. Nett,[1,4] Madeleine Bieg,[2,4] and Melanie M. Keller[3]

[1]Empirical Educational Research, Augsburg University, Germany, [2]Department of Empirical Educational Research, University of Konstanz, Germany, [3]Department of Physics Education, Leibniz Institute for Science and Mathematics Education, Germany, [4]Ulrike Nett and Madeleine Bieg contributed equally to this work

Abstract. Although the popularity of research on academic emotions is on the rise, little is known about the extent to which these emotional experiences are due to stable (trait) versus situational (state) influences. In the present paper, we applied the latent state-trait approach (LST) to multiple state assessments of five frequently experienced discrete academic emotions (enjoyment, pride, anger, anxiety, boredom) to disentangle their trait versus state variance components. We had two main aims: (1) to identify the differential contributions of the person-specific (trait) and situation-specific (state) variance components of discrete academic emotions, and (2) to examine the relations between different discrete academic emotions with regard to their latent trait and latent state residual components. Eight hundred thirty-seven German students participated in this diary study that lasted 2–3 weeks. During this time, students responded to short (two items per emotion) questionnaires asking about their lesson-specific state emotions in mathematics. The results revealed that for each academic emotion the trait variance and state residual components were of about equal size. Further, while differently valenced (positive vs. negative) latent *trait* components of students' emotions were mostly uncorrelated (with the exception of boredom), differently valenced latent *state* residual components of students' emotions were negatively correlated. We discuss our findings in relation to the structure of current affect and highlight their implications for classroom practices.

Keywords: academic emotions, state emotions, trait emotions, latent state-trait approach

"Maya is happy in math class, but William is anxious." Statements like these may be common in everyday conversations to describe students. However, can we really talk about a student as being happy or anxious in terms of a stable trait? Or are emotions instead highly fluctuating and situation-specific making it impossible to render any conclusions about their stability?

In psychological research, a defining characteristic of emotions – more so than other constructs – is their variability over time (Frijda, 2007). Emotional experiences, however, are also assumed to have stable aspects. The variability in an emotion would therefore describe deviations from a "baseline" or stable level of that emotion (e.g., Spielberger, Gorsuch, & Lushene, 1970). Determining the stability and variability in emotional experiences, and investigating how these stable and variable parts of various discrete emotions are interrelated, is a worthwhile and necessary endeavor to advance our knowledge regarding the very nature of emotions.

In recent years, education researchers have taken an ever-increasing interest in students' emotions during learning and achievement-related situations (Pekrun & Linnenbrink-Garcia, 2014). Academic emotions are proposed to consist of affective (core feeling), expressive (e.g., facial expression), cognitive (thoughts), physiological (e.g., heart rate), and motivational (fight or flight) components (e.g., Frenzel & Stephens, 2013; Kleinginna & Kleinginna, 1981). Furthermore, academic emotions can be categorized on a two-dimensional taxonomy consisting of valence and activation (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011; Feldman Barrett & Russell, 1998). The inclusion of the activation dimension (e.g., activating vs. deactivating emotions) in this taxonomy renders it possible to classify these emotions specifically as discrete emotions (e.g., positive activating: enjoyment, pride, positive deactivating: relief, negative activating: anger, anxiety, negative deactivating: boredom) instead of classifying them generally as positive and negative affect.

Researchers typically assess academic emotions with trait emotion questionnaires that ask students to report about their emotions “in general” regarding a specific school subject (e.g., math, foreign language, etc.; Pekrun et al., 2011; Pekrun, Goetz, Titz, & Perry, 2002a; Zeidner, 2007). This method of assessment implies that emotions are to some extent stable constructs or *traits*. Educational researchers, however, have recently begun to ask students about their *state* academic emotions – that is, academic emotions as experienced in the moment – and thus take into account the fluctuation of emotions from situation to situation or the interaction between the person and situation (Ahmed, van der Werf, Minnaert, & Kuyper, 2010; Becker, Goetz, Morger, & Ranellucci, 2014; Bieg, Goetz, & Hubbard, 2013; Goetz, Lüdtke, Nett, Keller, & Lipnevich, 2013; Goetz et al., 2014; Keller, Chang, Becker, Goetz, & Frenzel, 2014; Nett, Goetz, & Hall, 2011).

State-trait theories of emotions (e.g., Nesselroade, 1988) propose that a person’s current emotion level is determined by both stable (i.e., trait) and variable (i.e., situation-specific circumstances) factors. It is still unknown, however, how these two components of emotions – trait and state – overlap (i.e., how much trait is in state), especially in academic contexts, and whether and to what extent one can infer one from the other. Thus, a comprehensive analysis of trait and state components of academic emotions is needed to determine the stability and variability of students’ emotional experiences. Furthermore, differences in the stability and variability of discrete emotions would further expand our knowledge about whether various discrete emotions can be considered differentially trait-like or variable.

Disentangling the person-specific (trait) and situation-specific (state) variance components of academic emotions can be accomplished methodologically using the latent state-trait approach (LST; Steyer, Ferring, & Schmitt, 1992; Steyer, Mayer, Geiser, & Cole, 2015), which separates measurement error, trait-specific (stable) and variable (situation-specific; termed latent state residuals) variance components of emotions. In the present study we applied LST analysis to multiple state measures of students’ self-reported discrete emotions (enjoyment, pride, anger, anxiety, boredom). We sought to determine the extent to which students’ academic emotions are more “trait-like” or “state-like” by separating the stable and variable variance components of these emotions. Furthermore, we aimed to analyze the relations between latent trait components and latent state residuals of various discrete emotions. Examining these relations should provide insight into the interrelatedness of emotional traits and the covariation of academic emotions in a specific situation while taking into account the person’s trait level of emotion.

Stability and Variability of Emotions

The study of stable dispositions or traits has a long tradition in psychology (Carr & Kingsbury, 1938; Cattell, 1946). Indeed, the entire field of personality psychology is dedicated to the study and assessment of stable personality traits, which includes examining interindividual differences in these traits. Knowing how individuals usually behave (i.e., individuals’ reaction tendencies) allows us to predict their future behavior to some extent. Beyond well-researched dispositions, such as neuroticism, extraversion, and conscientiousness, personality researchers also study emotions and conceptualize them as stable dispositions (Epstein, 1977; Izard, Libero, Putnam, & Haynes, 1993).

A prominent example of state-trait theory is Spielberger’s conceptualization of trait anxiety and trait anger (Spielberger et al., 1970; see also Mischel, 1969; Zuckerman, 1976). This approach suggests that individuals have dispositions to experience certain emotions or tendencies to react with certain emotions across situations. For example, a highly trait anxious person should react more anxiously in different situations than a low trait anxious person. Furthermore, individuals’ emotional experiences are assumed to fluctuate around a person-specific trait level of the respective emotion. In general, traits can be defined as stable characteristics or individuals’ dispositions that predispose them “to perceive situations in particular ways and to react in a consistent manner in a wide variety of situations” (Zuckerman, 1976, p. 133). Trait emotions are most commonly assessed by means of questionnaires that ask individuals to estimate their emotional experience *in general* (Pekrun et al., 2011; Watson, Clark, & Tellegen, 1988; Zeidner, 2007).

In addition to the conceptualization of emotions as stable traits, another perspective suggests that emotions – perhaps even more so than other psychological constructs – are highly situation-specific, variable, and malleable (Ekman, 1992; Spielberger, 1972). This perspective is, for example, reflected in research that tries to experimentally manipulate emotions irrespective of individuals’ trait characteristics (Parkinson, 1985; Polivy, 1981). Furthermore, intraindividual variability in emotions (or affect as a broader construct) is an intensively researched area in emotion research suggesting that the fluctuations of emotions need to be considered to understand individuals’ emotional lives (e.g., Eid & Diener, 1999; Röcke & Brose, 2013). One approach often used to assess situation-specific emotions is to ask individuals about their emotions *right now* in a specific situation. This method enables researchers to investigate the variability of emotions when assessed across several measurement points (e.g., experience-sampling or diary methods; see Bolger, Davis, & Rafaeli, 2003). This approach may help researchers attain more direct and realistic insight into individuals’ actual emotional experiences. It should be noted, however, that a single state emotion assessment is still assumed to be influenced by the person’s emotional trait and the interaction between the person and the situation (Steyer et al., 1992). In other words, various state emotional assessments should share a common variance that can be attributed to the individual’s emotional trait.

Only a few studies to date have directly examined the extent to which emotions (general or academic) are stable versus variable. Some studies that have used multilevel analyses on state emotions have reported that a significant proportion of the variance is attributable to the individual’s emotional trait (Ahmed et al., 2010; Goetz, Lüdtke, et al., 2013; Merz & Roesch, 2011). Eid (1997) measured the consistency (trait) and specificity (state) of two positive (satisfied, happy) and two negative (dissatisfied, unhappy) mood states on four occasions and found higher specificity (state) coefficients for positive than for negative mood states. He further concluded that personal (trait) as well as situational (state) and/or interactional effects influence the respective mood states. In a similar vein, Yasuda and colleagues found that 56.9% to 63.5% of the reliable variance in positive affect and 48.2% to 60.6% of the reliable variance in negative affect were attributable to state-dependent or intraindividual variability when using a LST approach (Yasuda, Lawrenz, van Whitlock, Lubin, & Lei, 2004). It remains open to question, however, to what extent *discrete* academic emotions are stable versus variable. To the best of our knowledge, the present study is the first to delve into this question.

Relations Between Variable and Stable Components of Discrete Emotions

Students' emotional landscapes are highly complex, one reason being that they usually experience discrete yet interrelated emotions such as enjoyment, pride, anxiety, and anger while in their classrooms. Therefore, in addition to assessing the amount of stability and variability in emotions, it is also important to assess how these different emotions interrelate.

Bearing in mind the separation of emotional experiences into stable traits and situation-dependent states, the interrelation of discrete emotions can be addressed in two ways. First, one can examine how discrete emotions relate at the trait level; that is, does the trait level of a certain emotion covary with the trait level of another emotion? For example, does the level of trait anxiety relate to the level of trait anger or trait enjoyment? Second, at the state level, one can examine whether and to what extent the situation-specific (state) deviations from the latent trait of discrete emotions covary within a given situation. For example, in a certain situation does the deviation from the trait anxiety level differ from the deviation of the trait anger level?

With regard to the first question, education researchers have conducted several studies to examine the relationship between trait-level emotions as assessed by questionnaires. In general, significant medium-sized positive correlations have been reported for pairs of similarly valenced emotions (i.e., positive-positive emotion pairs such as enjoyment and pride, or negative-negative emotion pairs such as anxiety and anger) and significant negative correlations (low to medium in size) between differently valenced emotions (positive-negative emotion pairs; Goetz et al., 2012; Pekrun et al., 2011). In a few studies, researchers analyzed between-person correlations (which were interpreted as trait correlations) with experience-sampling data and found relatively weak between-person correlations of differently valenced emotions (Schimmack, 2003; Vansteelandt, Van Mechelen, & Nezlek, 2005; Zelenski & Larsen, 2000).

Regarding the second question, experience-sampling and diary studies have shed some light onto whether emotions covary in the same situation. Positive within-person correlations (to be interpreted as state correlations) were found for similarly valenced emotions and negative correlations were reported for differently valenced emotions (Schimmack, 2003; Zelenski & Larsen, 2000). This would imply that emotions of the same valence can be experienced simultaneously (a phenomenon called emotional blending; e.g., Vansteelandt et al., 2005), but that emotions of different valence cannot be experienced simultaneously.

The question of how positive and negative emotions interrelate has led to a hot debate that began almost 20 years ago (see Feldman Barrett & Russell, 1998, 1999) and is still going strong. This debate revolves around the question whether the valence of emotions is bipolar, meaning that differently valenced emotions are located on opposite ends of one dimension and are mutually exclusive in a certain situation (as would be indicated by a strong negative correlation between differently valenced emotions), or whether differently valenced emotions are independent of each other (as would be indicated by a close-to-zero correlation between differently valenced emotions; see Watson et al., 1988). Although a detailed discussion of the literature on the bipolarity of affect is beyond the scope of this article (see e.g., Russell & Carroll, 1999 for details on the correlations between momentary and extended affect), it is possible that this unresolved debate is in part due to a lack of differentiation regarding the role of stable (trait-related) and varying (state) components as they manifest in a certain emotional experience in a certain situation. It is important to point out here that all of the studies reported above used different analysis techniques of varied complexity (e.g., correlations: Zelenski & Larsen, 2000; or multivariate multilevel random coefficient models: Vansteelandt et al., 2005); none of them, however, were able to comprehensively disentangle the interrelations of stable and variable parts of emotional self-report data. In contrast, the LST approach used in the present research allows researchers to disentangle these two variance components of emotion self-reports and therefore has the potential to provide new and necessary insights to advance research also with regard to interrelations of these components.

Education researchers to date have examined the relationships between emotions as assessed by either classic trait self-report measures or aggregated or manifest state measures. A major drawback of trait emotion self-reports is that they give only limited insight into *actual* emotion experiences (e.g., memory bias; Robinson & Clore, 2002), and a major drawback of state emotion self-reports is that they are situation-specific and influenced by stable personality characteristics (e.g., Spielberger, 1972). The present research therefore aimed to investigate the interrelations of latent traits versus the interrelations of latent state residuals of discrete emotions with the LST approach. Thus, we were able to simultaneously answer questions such as: "Are we able to infer a person's trait regarding a specific emotion when knowing their other emotional traits?" and "Do deviations from two different trait emotions covary in the same situation?".

The Present Research

Assessing emotions "as they are experienced" is a prominent development in recent years (Hektner, Schmidt, & Csikszentmihalyi, 2007). Following up on the idea that there is a trait-specific component in each state emotion experience (Spielberger, 1972), one aim of the present study was to "extract" the trait component of students' emotions from multiple state emotion assessments. The LST approach is therefore key to achieving this aim because this method can disentangle the trait-specific and state-specific variance components of discrete emotions at specific measurement points (Steyer, Schmitt, & Eid, 1999; Steyer, Schwenkmezger, & Auer, 1990; Steyer et al., 2015). LST models have been mainly developed and applied to research questions examining the differential effects of stable personality traits as compared to situation or person × situation effects and random measurement error. In this approach, each measurement point is decomposed into a latent part (latent state) and error variance. This latent part is further decomposed into a latent trait component and a latent state residual. Thus, the latent trait component represents a stable component of an emotion experience across situations, whereas the latent state residual is the situation-specific deviation from the trait that results from situation-specific influences and/or interactions between the individual and the situation.

Using this approach, the reliability, consistency, and specificity of a construct can be assessed. Reliability in this sense refers to the proportion of variance that is not due to measurement error, but rather the variance that is explained by the model, namely by the latent state residual and the latent trait component. Consistency is the proportion of variance explained by the latent trait component, and specificity is the proportion of variance explained by the latent state residual. Consistency and specificity thus add up to reliability (Geiser & Lockhart, 2012). The LST approach therefore allowed us to determine the stability and variability of students' academic emotions. Extracting a trait component from several states is different from previous applications of LST that used multiple trait assessments to separate latent traits from state residuals (Geiser & Lockhart, 2012). In fact, few studies have used the LST approach with state emotion data (Eid, 1997; Eid & Diener, 1999).

In addition to determining stability and variability in academic emotions, we also investigated the correlations between five discrete emotions: enjoyment, pride, anxiety, anger, and boredom. These frequently occurring academic emotions (Pekrun et al., 2011) have been found to correlate at the trait and state level in several studies (Bieg et al., 2013; Pekrun et al., 2011). In these studies, however, the reported trait correlations came from classic trait assessments (assumed to be biased due to e.g., subjective beliefs such as self-concept; Bieg, Goetz, & Lipnevich, 2014; Robinson & Clore, 2002), whereas the reported state correlations came from correlations of manifest states without considering the stable (trait) versus situation-specific components of state emotions. It is therefore still open to question how these emotions relate to each other when using latent trait components (reflecting the stable component) versus latent state residuals (reflecting the situation-specific/variable component) as extracted by LST from various state measures. As mentioned earlier, this is a worthwhile endeavor because it would allow us to determine, for example, whether an anxious student can also be a happy student (trait level) and to what extent discrete emotions covary in the same situation (correlation of state residuals) taking into account a person's individual trait level of these emotions – questions that are answerable using the LST approach.

In the present diary study we investigated five discrete emotions students experience specifically in mathematics lessons. Two criteria guided the choice of emotions for the study. First, we included emotions that students frequently experience in the classroom (Pekrun et al., 2002a). Second, we included emotions that came from various categories based on the valence/activation taxonomy. These emotions included two frequently occurring positively valenced activating emotions, enjoyment and pride; two frequently occurring negatively valenced activating emotions, anxiety and anger; and one (slightly) negatively valenced deactivating emotion, boredom. We included boredom because students experience this emotion rather frequently in the academic context (e.g., Nett et al., 2011). We decided not to include positive deactivating emotions (e.g., relief, relaxation after success) because these emotions appear to be less relevant during instruction (i.e., class-related academic emotions) and because findings about these emotions have been inconclusive (Pekrun, Goetz, Titz, & Perry, 2002b). Similar to previous studies (e.g., Goetz, Lüdtke, et al., 2013; Goetz et al., 2014), the emotions in the current study were assessed by items addressing the specific emotions explicitly (see Appendix).

Our aims for the present study using the LST approach were twofold. First, we aimed to identify the extent to which the experience of discrete academic emotions is stable versus variable. We hypothesized that each discrete emotion would have substantial person-specific (trait) as well as situation-specific (state) variance. Second, we aimed to examine the relationships between different discrete emotions. In particular, we examined (a) the relationships between the latent trait components as extracted by the LST from the state assessments of discrete emotions and (b) the relationships between the latent state residuals of discrete emotions.

Materials and Method

Participants and Procedure

Eight hundred thirty-six students (54.30% female) from a total of 43 9th and 10th grade classes (M_{age} = 15.58; SD_{age} = 0.71) in the highest-achieving track of the three-tiered German school system (Gymnasium) participated in our diary study. Over a two to three week time period, immediately after each mathematics lesson, students filled out a short questionnaire contained within a booklet. Depending on their curriculum, students attended on average 7.05

Table 1. Means and standard deviations of emotion measures

	Occasion 1		Occasion 2		Occasion 3		Occasion 4		Occasion 5		Occasion 6	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Enjoyment 1	2.90	1.15	2.96	1.20	2.83	1.14	2.80	1.16	2.89	1.14	2.90	1.25
Enjoyment 2	2.78	1.16	2.95	1.23	2.75	1.21	2.81	1.18	2.80	1.16	2.81	1.24
Pride 1	2.51	1.10	2.74	1.15	2.67	1.12	2.67	1.13	2.72	1.10	2.67	1.21
Pride 2	2.58	1.19	2.78	1.21	2.68	1.19	2.73	1.16	2.77	1.16	2.71	1.21
Anxiety 1	1.95	1.07	1.90	1.05	1.97	1.05	1.88	0.97	1.84	0.99	1.84	1.03
Anxiety 2	1.33	0.77	1.39	0.80	1.45	0.86	1.44	0.83	1.40	0.80	1.49	0.89
Anger 1	1.42	0.82	1.49	0.90	1.47	0.80	1.48	0.87	1.45	0.83	1.49	0.89
Anger 2	1.74	1.11	1.80	1.12	1.78	1.04	1.77	1.08	1.70	1.01	1.79	1.09
Boredom 1	2.44	1.16	2.41	1.19	2.45	1.17	2.41	1.19	2.35	1.15	2.32	1.25
Boredom 2	2.48	1.11	2.43	1.18	2.52	1.14	2.41	1.17	2.35	1.14	2.37	1.25

(*SD* = 2.34) mathematics lessons during this period of time. Taking into account dropout rates and model complexity, we decided to include six measurement points in our models. The first mathematics lesson was rated by 93.20% of the students (all 43 classes participated); subsequent lessons were rated with decreasing frequency. The sixth lesson was rated by 73.24% of the students (five entire classes were missing). After the 6th lesson, nine or more entire classes were missing (less than 67.50% of the students), thus we excluded these lessons from the present analyses. Students were assured of the anonymity of the assessment and that no personal identifiers would be retained in the final dataset. Participation in this study was voluntary and students could terminate their participation at any time.

Measures

Within the short diary questionnaire, the five discrete emotions – enjoyment, pride, anxiety, anger, and boredom – were measured by two items each. All of the items were adapted from the class-related emotions scale from the Achievement Emotions Questionnaire (AEQ) by Pekrun et al. (2011) and referred to the mathematics lesson students had just attended (i.e., after the situation). The items were introduced with the wording: "In this lesson..." and ended with an emotion-specific part (e.g., "...I felt anxious."). All items were assessed on a 5-point Likert scale ranging from 1 (= *strongly disagree*) to 5 (= *strongly agree*). Means and standard deviations for the 10 emotion items at each of the six occasions (i.e., the mathematics lessons) are depicted in Table 1. The exact wording of the items in German and their English translations are included in the Appendix, Table A1.

Data Analysis

We addressed our research questions by conducting a series of LST models. In the first step, we calculated a LST model for each emotion (enjoyment, pride, anxiety, anger, and boredom) with indicator-specific trait factors (Geiser & Lockhart, 2012) that resulted in six state factors (one state factor for each occasion) indicated by two items each, and two trait factors indicated by the same item measured at six different occasions. The two indicator-specific trait factors were correlated (see Figure 1). All models were tested for strong measurement invariance. Strict measurement invariance was not assumed. From these models, reliability, consistency, and specificity indicators (Geiser & Lockhart, 2012; Steyer et al., 1999) were derived for each emotion. As mentioned previously, reliability indicates the degree to which the observed interindividual differences can be explained by reliable sources of variance; consistency indicates the degree to which variance is due to stable person-specific traits; specificity indicates the degree to which variance is due to the situation or the interaction between person and situation. Consistency and specificity sum up to reliability (Geiser & Lockhart, 2012).

To analyze the present data, we could have used other statistical models, such as autoregressive models (e.g., the STARTS model; Kenny & Zautra, 2001), that take into account changes in emotional experiences over a period of time. Because we did not assume any systematic changes in the experiences of students' emotions during the assessment period, we preferred to focus on the separation of stable latent traits and latent state residuals. Moreover, autoregressive models usually require equally spaced assessments (Biesanz, 2012), a condition that was not met in our emotion assessment.

In the second step, we tested a series of ten models that combined pairwise (i.e., two at a time) the LST models for the five emotions (Figure 2) thereby adapting the "indicator-specific trait model for two methods" according to Geiser and Lockhart (2012, p. 273). Because the graphical representation of these models resembles a butterfly, we will refer to these models as "butterfly models of emotions." In these models, all indicator-specific latent trait factors were correlated with each other, resulting in four correlational values between the two emotional traits for each emotion. Furthermore, the state residual factors pertaining to the different emotions that were measured at the same measurement point were correlated with each other, resulting in six correlations between latent state residuals of the six measurement occasions (see Figure 2). The correlations between the latent trait factors represent the relations of the stable components in students' emotions (i.e., part (a) of our second research question), whereas the

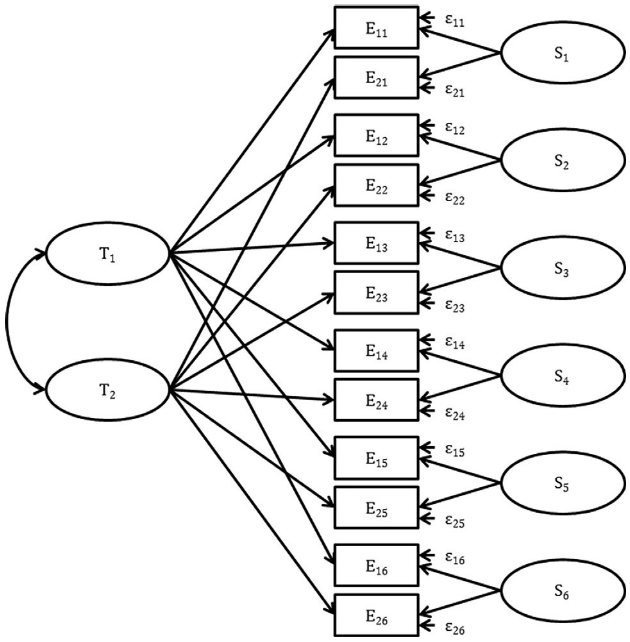

Figure 1. Indicator-specific LST model with two correlated latent trait variables. In the presented final models, strong measurement invariance was assumed, thus all factor loadings were fixed to equal 1 and no change over time of the intercepts was assumed. E_{ij} refers to the emotion-specific item *i* at occasion *j*. T_i refers to the latent trait for emotion specific item *i*. S_j refers to the latent state residual for occasion *j*.

correlations between the latent state residuals indicate the co-occurrence of the two emotions (i.e., part (b) of our second research question; Geiser & Lockhart, 2012). In order to be able to summarize the four trait correlations and the six state residual correlations for each pair of emotions and thus gain a comprehensive picture of the interrelations of latent state residuals between different discrete emotions, we Fisher-*z*-transformed, averaged, and retransformed the four latent trait and six latent state residual correlation scores. These retransformed values can be interpreted as mean correlations of the latent traits and latent state residuals.

All of the analyses were conducted with M*plus* software version 7.11 (Muthén & Muthén, 1998–2013). The nested data structure was taken into account using the TYPE IS COMPLEX command in M*plus* and applying the maximum likelihood robust (MLR) estimator. Missing data was accounted for by using the full information maximum likelihood estimator (FIML).

Results

Trait- and State-Specific Contributions to Academic Emotions

In order to determine the trait- and state-specific variance of academic emotions, we evaluated indicator-specific LST models. Models were tested for strong measurement invariance. Because the model fits of the models with strong measurement invariance were equal to or only slightly worse than the models with weak invariance, and because they provided very good fit (Table 2) based on fit indices such as the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the root mean squared error of approximation (RMSEA), and the standardized root mean squared residual (SRMR; Hu & Bentler, 1999), strong measurement invariance was assumed. Alternative models, namely single-trait multistate models (Geiser & Lockhart, 2012), provided worse model fit indices than the proposed models across all emotions.

The variances of latent traits and latent state residuals were meaningful for each emotion. Indicator-specific latent traits were highly correlated with one another for each emotion (Table 3). Table 4 provides summaries of the item-specific reliability, consistency, and specificity scores; Appendix Table A2 depicts the complete item-specific reliability, consistency, and specificity scores. Across all emotion items, reliability ranged from .52 (for one of the anxiety items) to .84 (for one of the boredom items), indicating that the applied items were sufficiently reliable. Consistency ranged from .26 (for one of the anger items) to .42 (for one enjoyment and one boredom item). These consistency values represent the proportion of the item's *total* variance that can be explained by stable traits. When consistency was calculated referring to the "true" (measurement-error free) variance of an item (i.e., based on item reliability), between 37.88% and 68.07% of the variance was explained by stable traits.

The specificity scores ranged from .17 (for one of the anxiety items) to .50 (for one anger and one boredom item). These specificity values represent the proportion of the item's *total* variance that can be explained by situation-specific influences. Again, when consistency was calculated referring to the "true" (measurement-error free) variances, between 31.93% and 62.12% was due to situation-specific characteristics or due to the interaction between individual and situation.

Across all emotions, on average 48.68% of the explained variance was due to latent traits and 51.60% of the explained variance was due to situation-specific (state) characteristics. These results suggest that, as hypothesized, math-related classroom emotions incorporate a meaningful stable (trait) component as well as a meaningful situation-specific (state) component. Results further suggest that the stable versus situation-specific variance components are quite balanced for each emotion.

Relations Between Trait and State Academic Emotions

As mentioned above, in the "butterfly models of emotions" addressing interrelations between different discrete emotions (see Figure 2), strong measurement invariance was assumed for the LST models. Fit indices were all satisfactory according to Hu and Bentler (1999, see Table 5). For each pair of emotions, the indicator-specific

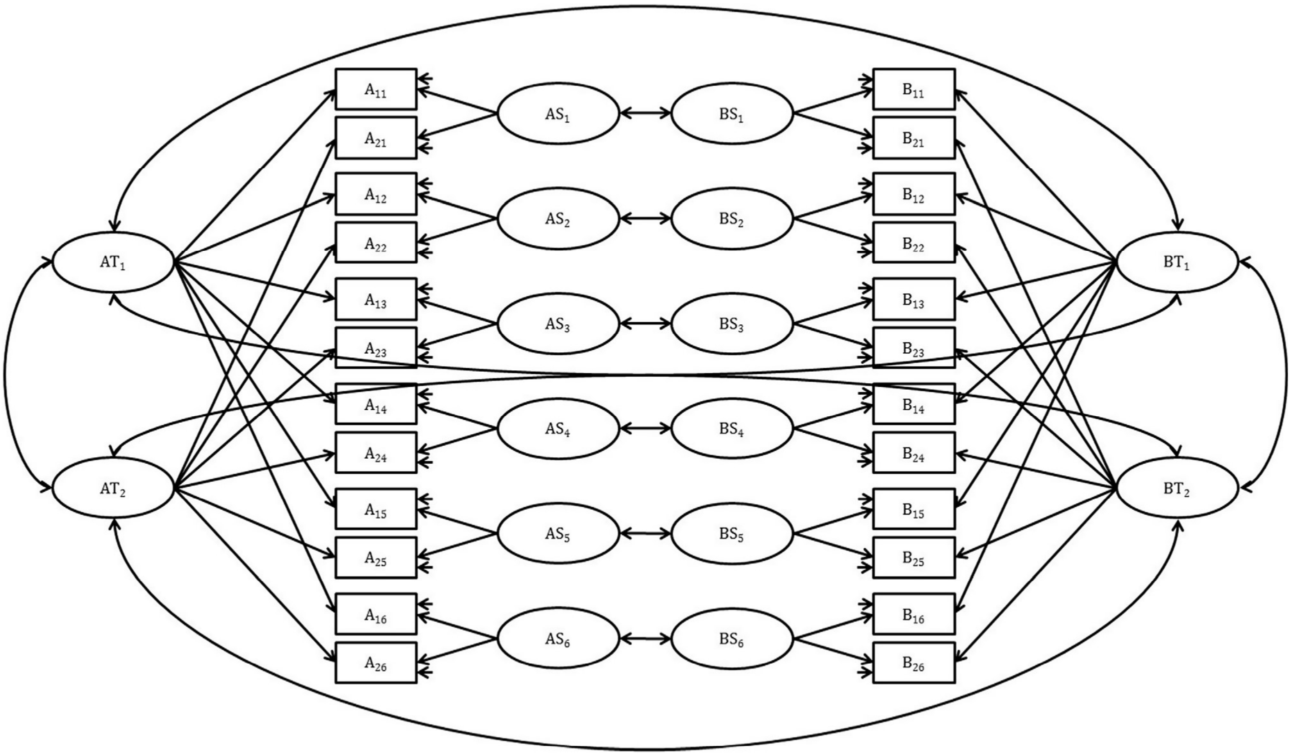

Figure 2. LST model with two emotions ("butterfly" model). In the calculated models, strong measurement invariance of the two included LST models was assumed. A_{ij} and B_{ij} refer to the emotion-specific item *i* at occasion *j* of emotions A and B. AT_i and BT_i refer to the latent trait for emotion-specific item *i* of the emotions A and B, AS_j and BS_j refer to the latent state residual for occasion *j* of the emotions A and B.

Table 2. Fit indices of indicator-specific latent state-trait models

	MI	Chi-Square; *df* (*p*)	CFI/TLI	RMSEA	SRMR
Enjoyment	Strong MI	87.73; 67 (.046)	0.99/0.99	0.02	0.05
	Weak MI	73.46; 57 (.070)	0.99/0.99	0.02	0.04
Pride	Strong MI	102.66; 67 (.003)	0.99/0.99	0.03	0.05
	Weak MI	85.18; 57 (.009)	0.99/0.99	0.02	0.05
Anxiety	Strong MI	116.97; 67 (.000)	0.97/0.97	0.03	0.05
	Weak MI	87.18; 57 (.006)	0.98/0.98	0.03	0.05
Anger	Strong MI	66.57; 67 (.492)	1.00/1.00	0.00	0.04
	Weak MI	61.28; 57 (.325)	1.00/1.00	0.01	0.04
Boredom	Strong MI	103.16; 67 (.003)	0.99/0.99	0.03	0.05
	Weak MI	95.80; 57 (.001)	0.99/0.99	0.03	0.05

Notes. Strong MI are models with strong measurement invariance assuming invariance of factor loadings and intercepts; Weak MI are models with factorial measurement invariance assuming invariance of factor loadings only.

latent traits were correlated with each other, resulting in four correlations (upper right part of Table 6). Further, the six latent state residuals for each emotion pair were correlated as well (lower left part of Table 6).

With regard to the latent trait correlations, the two positively valenced emotions, enjoyment and pride, were positively correlated (large effect). The two negatively valenced emotions, anxiety and anger, were also positively correlated (large effect). The correlations between differently valenced emotions, however, were negligible or of very small effect size. Trait-level boredom was negatively correlated with enjoyment (large effect), negatively correlated with pride (medium effect), positively correlated with anger (medium effect), and uncorrelated with anxiety. Overall, these results indicate that similarly valenced emotions (enjoyment and pride; anxiety and anger) are highly positively correlated at the trait level while differently valenced emotions – with the exception of boredom – are independent of each other at the trait level.

Table 3. Variances of latent traits and latent state residuals

	Enjoyment	Pride	Anxiety	Anger	Boredom
Var(S_1)	0.49	0.58	0.30	0.37	0.58
Var(S_2)	0.57	0.53	0.24	0.39	0.62
Var(S_3)	0.47	0.46	0.28	0.24	0.51
Var(S_4)	0.52	0.45	0.16	0.29	0.52
Var(S_5)	0.46	0.45	0.20	0.27	0.48
Var(S_6)	0.64	0.65	0.35	0.39	0.79
Var(T_1)	0.53	0.42	0.34	0.24	0.52
Var(T_2)	0.44	0.49	0.23	0.33	0.51
r_{T1T2} (p)	0.81 (< .001)	0.92 (< .001)	0.75 (< .001)	0.84 (< .001)	0.95 (< .001)

Notes. Variances are extracted from LST models as depicted in Figure 1. S_j refers to the latent state residual for occasion j. T_i refers to the latent trait for emotion specific item i.

Table 4. Summary of item-specific reliabilities, consistencies, and specificities

	Variance (ε_{it})	Reliability	Consistency	Specificity
Enjoyment				
Mean	0.39	0.72	0.35	0.37
Medium	0.39	0.73	0.34	0.37
Min	0.27	0.65	0.29	0.34
Max	0.51	0.79	0.42	0.43
Pride				
Mean	0.37	0.72	0.34	0.39
Medium	0.37	0.72	0.34	0.38
Min	0.31	0.69	0.29	0.33
Max	0.43	0.75	0.37	0.45
Anxiety				
Mean	0.33	0.63	0.33	0.30
Medium	0.35	0.60	0.33	0.29
Min	0.17	0.52	0.29	0.17
Max	0.55	0.76	0.37	0.46
Anger				
Mean	0.33	0.67	0.31	0.36
Medium	0.31	0.67	0.31	0.35
Min	0.13	0.54	0.26	0.23
Max	0.56	0.82	0.38	0.50
Boredom				
Mean	0.28	0.80	0.38	0.42
Medium	0.26	0.80	0.38	0.40
Min	0.24	0.75	0.32	0.37
Max	0.35	0.84	0.42	0.50

Notes. Reliability, Consistency, and Specificity were calculated according to Geiser and Lockhart (2012). As all factor loadings in the models were fixed to equal 1, the formulas can be simplified to the following:

$$\mathrm{Consistency}(\mathrm{E_{it}}) = \frac{\mathrm{Variance}(\mathrm{T_i})}{\mathrm{Variance}(\mathrm{E_{it}})}$$

$$\mathrm{Specificity}(\mathrm{E_{it}}) = \frac{\mathrm{Variance}(\mathrm{S_t})}{\mathrm{Variance}(\mathrm{E_{it}})}$$

$$\mathrm{Reliability}(\mathrm{E_{it}}) = 1 - \frac{\mathrm{Variance}(\varepsilon_{\mathrm{it}})}{\mathrm{Variance}(\mathrm{E_{it}})} = \mathrm{Consistency}(\mathrm{E_{it}}) + \mathrm{Specificity}(\mathrm{E_{it}})$$

Please note that the table includes summaries in order of means, mediums, minimum, and maximum scores. Thus the consistency and specificity scores in one line of this table do not sum up to the reliability scores. Single scores for each item are depicted in the Appendix Table A2.

Table 5. Fit indices of the "Butterfly Models of Emotions"

	Chi-square; *df* (*p*)	CFI/TLI	RMSEA	SRMR
Enjoyment-Pride	551.48, 268, (< .001)	0.96/0.96	0.04	0.05
Enjoyment-Anxiety	422.14, 268, (< .001)	0.97/0.97	0.03	0.05
Enjoyment-Anger	385.57, 268, (< .001)	0.98/0.98	0.02	0.04
Enjoyment-Boredom	369.06, 268, (< .001)	0.99/0.99	0.02	0.05
Pride-Anxiety	413.85, 268, (< .001)	0.97/0.97	0.03	0.05
Pride-Anger	358.79, 268, (< .001)	0.98/0.98	0.02	0.04
Pride-Boredom	345.49, 268, (.001)	0.99/0.99	0.02	0.05
Anxiety-Anger	406.40, 268, (< .001)	0.97/0.97	0.03	0.05
Anxiety-Boredom	391.51, 268, (< .001)	0.98/0.98	0.02	0.05
Anger-Boredom	350.18, 268, (< .001)	0.99/0.99	0.02	0.04

Notes. In these models, the number of parameters is more than the number of clusters minus the number of strata with more than one cluster, therefore the standard errors may not be fully trustworthy when using the command TYPE IS COMPLEX in M*plus*. However, the correlation values are equal to correlation values without taking the nested structure of the data into account, and *p*-values are very similar, but slightly more conservative, than the *p*-values without taking the nested structure of the data into account.

The relations between the latent state residuals of discrete emotions were surprisingly different from the latent trait correlations. Most correlations were significant (with the exception of one state residual correlation between pride and anxiety, most of the correlations between boredom and anger, and all of the correlations between boredom and anxiety). Again, the similarly valenced emotions (enjoyment and pride; anxiety and anger) were positively correlated with each other (large effects). In contrast to the respective latent trait correlations, differently valenced emotions were negatively correlated with one another (medium effects). Again, boredom exhibited its own distinct correlational pattern: Boredom was negatively correlated with enjoyment (large effect) and pride (medium effect), sometimes significantly correlated with anger (small effects), and again uncorrelated with anxiety. Overall, these results indicate that, when accounting for the person-specific (trait-level) components of these emotions, similarly valenced emotions (enjoyment and pride; anxiety and anger) covary in the situation, whereas differently valenced emotions do not covary.

Discussion

The aims of the present study were twofold. Our first goal was to use the LST approach to examine the differential contributions of the person-specific (trait) and situation-specific (state) variance components of five frequently experienced class-related emotions in mathematics. Our second goal was to explore the relations between different discrete emotions with regard to (a) their latent trait components as extracted by the LST from state assessments of discrete emotions and (b) their latent state residual components.

Trait- and State-specific Contributions to Academic Emotions

"Can we really talk about a student as being happy or anxious in terms of a stable personality characteristic? Or are emotions variable and situation-specific constructs that prohibit us from drawing any conclusions about their stability?" These questions about the stability and variability of emotions were asked at the beginning of the paper and one aim of the present study was to separate the trait-specific variance from the state-specific variance of state academic emotions. We analyzed state emotions because they are assumed to be more reliable measures of students' actual emotional experiences than classic trait self-reports (Bolger et al., 2003; Robinson & Clore, 2002; Goetz, Bieg, et al., 2013). The results showed that both trait and state components made meaningful contributions to the emotions students experienced in mathematics lessons. In other words, there were stable contributions of traits across situations that influence students' emotions but the situation also caused variability in students' state emotional experiences.

An evaluation of the consistency and specificity scores revealed that the variance that could be explained by systematic sources (reliability) was almost equally divided between trait- and state-specific variance for all discrete emotions. This suggests that research on the experience of academic emotions needs to take into account both sources of variance in academic emotions. Moreover, the specificity coefficients of discrete emotions in the present study were comparably higher than specificity coefficients found in previous studies that examined facets of state self-esteem or competitive anxiety in sports measured on two occasions (Hank, 2015; Ziegler, Ehrlenspiel, & Brand, 2009) – although Eid (1997) also found relatively high specificity coefficients for mood ratings on four measurement occasions. Because we know of no other studies using the LST approach to explicitly examine academic emotions, replications of our results are highly warranted.

An even more detailed evaluation of each emotion showed that, except for anxiety, specificity coefficients were slightly higher than consistency coefficients – a finding that is in line with the conceptualization of anxiety as stable trait (Spielberger et al., 1970). Nevertheless, further research should investigate if this finding is significant and stable across different measurement settings.

Table 6. Pairwise correlations between latent traits and occasion-specific state residuals

	Enjoyment		Pride		Anxiety		Anger		Boredom	
	r	(*p*)	*r*	(*p*)	*r*	(*p*)	*r*	(*p*)	*r*	(*p*)
Enjoyment			0.57	(< .001)	−0.14	(0.006)	−0.17	(0.001)	−0.62	(< .001)
			0.62	(< .001)	−0.09	(0.025)	−0.28	(< .001)	−0.56	(< .001)
			0.71	(< .001)	0.06	(0.221)	−0.01	(0.863)	−0.46	(< .001)
			0.84	(< .001)	0.06	(0.315)	−0.07	(0.176)	−0.44	(< .001)
Mean *r*			0.70		−0.03		−0.13		−0.52	
Pride	0.80	(< .001)			0.08	(0.120)	0.10	(0.040)	−0.29	(< .001)
	0.84	(< .001)			0.08	(0.120)	0.01	(0.900)	−0.31	(< .001)
	0.95	(< .001)			0.11	(0.027)	0.08	(0.126)	−0.32	(< .001)
	0.87	(< .001)			0.11	(0.021)	−0.01	(0.817)	−0.34	(< .001)
	0.95	(< .001)								
	0.90	(< .001)								
Mean *r*	0.90				0.10		0.05		−0.32	
Anxiety	−0.44	(< .001)	−0.15	(0.076)			0.75	(< .001)	0.12	(0.103)
	−0.44	(< .001)	−0.40	(< .001)			0.76	(< .001)	0.09	(0.213)
	−0.34	(< .001)	−0.26	(0.003)			0.76	(< .001)	0.13	(0.039)
	−0.62	(< .001)	−0.47	(< .001)			0.65	(< .001)	0.04	(0.513)
	−0.31	(0.001)	−0.31	(0.001)						
	−0.34	(< .001)	−0.20	(0.019)						
Mean *r*	−0.42		−0.30				0.73		0.10	
Anger	−0.56	(< .001)	−0.29	(< .001)	0.60	(< .001)			0.34	(< .001)
	−0.44	(< .001)	−0.39	(< .001)	0.80	(< .001)			0.26	(< .001)
	−0.40	(< .001)	−0.34	(< .001)	0.77	(< .001)			0.36	(< .001)
	−0.53	(< .001)	−0.40	(< .001)	0.99	(< .001)			0.26	(< .001)
	−0.41	(< .001)	−0.36	(< .001)	0.95	(< .001)				
	−0.43	(< .001)	−0.30	(< .001)	0.92	(< .001)				
Mean *r*	−0.46		−0.35		0.90				0.31	
Boredom	−0.61	(< .001)	−0.28	(0.002)	0.11	(0.147)	0.26	(< .001)		
	−0.61	(< .001)	−0.46	(< .001)	-0.03	(0.649)	0.11	(0.098)		
	−0.63	(< .001)	−0.47	(< .001)	0.05	(0.503)	0.22	(0.003)		
	−0.55	(< .001)	−0.51	(< .001)	0.14	(0.270)	0.23	(0.028)		
	−0.50	(< .001)	−0.44	(< .001)	0.08	(0.319)	0.15	(0.059)		
	−0.54	(< .001)	−0.53	(< .001)	0.01	(0.909)	0.14	(0.188)		
Mean *r*	−0.58		−0.45		0.06		0.19			

Notes. Latent trait correlations are depicted above the diagonal, and latent state residual correlations are depicted below the diagonal. Mean *r*'s were calculated by Fisher-*z*-transforming the correlation values, taking the mean and retransforming the mean of the Fisher-*z*-transformed correlation values.

The present research focused specifically on class-related emotions in mathematics given that the experience of academic emotions is domain-specific (Goetz, Frenzel, Pekrun, & Hall, 2006; Goetz, Pekrun, Hall, & Haag, 2006). Thus, the measurements took place in rather homogeneous settings. This should be kept in mind when interpreting our findings, as the proportion of trait and state variance might differ based on the research setting (e.g., school situation or subject). For example, if we had examined several different math-related school situations – while doing homework, while studying, while taking tests, *and* while in class – we might have found a larger proportion of variance attributable to state influences than we found in the present study.

Further, the results imply the need to carefully distinguish between trait and state components when discussing emotions. Classic approaches, such as generalized trait questionnaires and aggregated state measures, cannot differentiate trait from state components, therefore any conclusions drawn from these approaches may be a mixture of conclusions about trait and state contributions. Thus, the LST approach offers a valuable and applicable methodological tool to separate stable (trait) from situation-specific (state) contributions.

Relationships Between Academic Emotions

"Can a happy student also be an anxious student?" In response to this question, our analyses suggest that the latent trait components of differently valenced emotions are virtually unrelated (except for the correlations between positive emotions and boredom that will be discussed below). Thus, based on our findings, knowing that a student is generally a happy person does not allow any conclusions to be made about whether this student is also generally an anxious person; these two emotional traits appear to be independent of each other. These results differ from prior findings on the correlations between manifest traits of differently valenced emotions in educational contexts (e.g., Pekrun et al., 2011) and once again highlight the need to separate trait- from state-specific variance components. Our results are therefore in line with previous studies that have found differently valenced emotions to be independent of each other rather than falling at opposite ends of a bipolar continuum (Feldman Barrett & Russell, 1998).

On the other hand, our findings suggest that the latent trait components of similarly valenced emotions (i.e., enjoyment and pride; anxiety and anger) are strongly positively related. This is in line with previous research on affect that distinguishes between positive and negative affect without taking into account discrete emotions (Watson et al., 1988).

The findings for trait-level boredom differed considerably from the other negative emotions included in the present study. The latent trait component of boredom was negatively related to both enjoyment and pride and positively related to anger but unrelated to anxiety. This correlational pattern is consistent with previous findings from boredom research that detail the complexity and ambiguity of this emotion (Goetz et al., 2014). Had boredom been excluded from our study, our findings could have led to the erroneous conclusion that differently valenced emotions are independent of each other, whereas similarly valenced emotions exhibit substantial correlations representing a general positive/negative emotional trait. In fact, this is the common finding in studies that do not include "ambiguous" emotions like boredom (Feldman Barrett & Russell, 1998). Interpreting the whole of our own findings – including boredom – leads to conclusions that are not uniform or clear-cut. Overall, the pattern of correlations based on all five emotions supports rather than refutes the existence of discrete academic emotions even at the trait level (Pekrun et al., 2011). This observation further highlights the importance of including additional discrete emotions in future studies (e.g., hope, shame, envy, etc.) to be able to draw a comprehensive picture of students' emotional landscapes in academic settings.

Turning to the state residuals of academic emotions, our final question concerned the covariation between academic state emotions when accounting for the trait-level emotions. In general, the results showed that the deviations from the trait levels of differently valenced emotions were usually negatively correlated with each other indicating mutual exclusivity, whereas similarly valenced emotions were positively correlated with each other indicating co-occurrence. The extremely high positive correlations between similarly valenced emotions (excluding boredom) could also imply that within certain academic settings, such as during a math lesson, enjoyment and pride and likewise anxiety and anger are experienced almost as one and the same emotion. These findings could, however, be very different in situations that induce more variability in certain emotions (e.g., during exams, levels of anxiety should be higher than during class instruction; during academic competitions, levels of pride should be higher than during class instruction). Thus, within situations that induce an extreme level of a certain emotion, it might be easier to identify differences in the experience of emotions, such as anxiety and anger or enjoyment and pride, than in typical classroom situations. In summary, these findings are in line with studies showing that similarly valenced emotions coexist within situations (Vansteelandt et al., 2005; Zelenski & Larsen, 2000), although these studies did not explicitly separate state from trait influences.

It is important to note here that Vansteelandt and colleagues' (2005) findings were derived from multivariate multilevel random coefficient modeling. In this type of analysis, for each person, each emotion at a specific measurement point is predicted by an overall intercept, a person-specific disposition, an emotional person-specific state, and an error term. While Vansteelandt and colleagues' (2005) approach and the LST approach in the present study can be used to answer similar research questions, two key methodological differences should be mentioned. First, the multivariate multilevel random coefficient model (Vansteelandt et al., 2005) estimates across individuals the average of the intraindividual correlations between two constructs (e.g., between two emotions: anxiety and anger). In such analyses, all measurement occasions are considered simultaneously within each individual. In contrast, the LST model used in the present study estimates the correlations of the latent state residuals separately for each measurement occasion, the result of which is each individual's deviation from his or her stable trait. Thus, the presented findings provide information about the situational impact of the measurement occasions included in the study. Second, whereas the multivariate multilevel random coefficient model enables researchers to analyze interindividual differences of within-person correlations (i.e., do certain pairs of emotions correlate more strongly for some individuals than for others?), the findings from the LST models herein provide between-person correlations of latent state residuals, which can be directly related to the between-person correlations of the latent traits. The different correlational patterns of latent trait emotions versus latent state residuals as found in the present study again highlight the importance of differentiating between trait versus state influences when analyzing academic emotions – an outcome that can be accomplished using the presented LST approach.

Limitations of the Present Study

Several limitations to our study need to be considered. For instance, our sample was restricted to the highest-achieving track of the German school system (Gymnasium) and to 9th and 10th graders. Thus, at present, we cannot provide any information about the generalizability of our results across age or achievement level. It is also important to consider whether a personality trait can be measured within a period of two to three weeks. For instance, it could be that the stable "trait" we measured during this period of time was actually a state that could have changed when, for instance, a new topic was introduced in the curriculum. This possibility seems unlikely, however, because the data was collected during lessons that likely spanned several topics. Moreover, because of the full measurement invariance we are confident that we assessed at least the current trait level of students' emotional experiences. Nevertheless, a much longer longitudinal study would be necessary to uncover systematic changes in trait levels of emotions.

In line with several other studies (e.g., Tsai, Kunter, Lüdtke, Trautwein, & Ryan, 2008; Vansteelandt et al., 2005), we assessed *state* emotions by asking students to evaluate their emotions at the *end* of a lesson; however, there are also studies that measure state emotions by asking students to evaluate their emotional experiences *right now* (i.e., directly after or during an emotion is experienced). Although it may appear as if studies that assess "right now" emotions are the only studies truly assessing state emotions, it is open to question whether a state emotion is only what a person experiences at one particular moment. In other words, how long is the retrospective time frame allowed to be before a state emotion "becomes" a trait emotion? This leads to the broader question of where the fine line is between "traits" and "states." Our own assessment of state emotions incorporates short retrospection and thus does not refer to an emotion experienced right now. Nevertheless, our assessment is still more episodic-based than asking about emotions "in general." A fruitful venue for future research should therefore include the investigation of trait- and state-specific variance components with more directly assessed emotions (e.g., experience sampling).

Conclusions and Future Prospects

State academic emotions consist of stable (trait) as well as situation-specific (state) components. In the present study, the LST approach offered a highly valuable approach to differentiate trait and state components of emotions and thus disentangle stability from variability in state emotional self-reports. Although researchers have long speculated about the importance of trait and state components of emotions, the methodological approaches to test these assumptions were restricted. Given our intriguing findings on the trait and state components of students' academic emotions and the correlations between different discrete trait and latent state residuals of emotions, we would like to encourage other researchers to also apply the LST approach in future studies to gain more insight into academic emotions. One valuable avenue for future research would be to investigate latent trait components in emotions as extracted from LST models and compare these latent trait components to trait emotions assessed by classic trait emotion questionnaires (e.g., AEQ; Pekrun et al., 2011). This comparison would allow for conclusions about the validity of students' judgments of their trait emotions in specific domains. The LST approach might be especially helpful when investigating certain situational influences on students' academic emotions (e.g., teaching methods). This should be even more so the case when researches aim to systematically vary situational conditions (e.g., experimentally manipulate teaching methods), because the LST approach can provide information about the impact of each specific situation on students' emotional experiences.

A question that frequently arises in relation to academic emotions is how to avoid negative and enhance positive emotions in the classroom (Goetz, Lüdtke, et al., 2013). When searching for antecedents of academic emotions and how to influence them, it is important to clearly differentiate between *dispositional* antecedents, such as neuroticism (e.g., Costa & McCrae, 1992), trait anxiety (e.g., Spielberger et al., 1970), or boredom proneness (e.g., Farmer & Sundberg, 1986), and *situational* antecedents such as instructional behavior (e.g., Daschmann, Goetz, & Stupnisky, 2011) or control and value appraisals (e.g., Goetz, Frenzel, Stoeger, & Hall, 2010). From this practical perspective, our results showing that roughly half of the variance in students' academic emotions is attributable to situation-specific factors imply that teachers can make a positive difference in their students' emotional lives.

Acknowledgments

We would like to thank Eva Becker und Thomas Goetz who were involved in study design and data collection. We would like to further thank Anne Frenzel, Philipp Döbler, Johannes Schult, and Fabio Sticca for helpful feedback regarding data analysis and conception of the present paper.

References

Ahmed, W., van der Werf, G., Minnaert, A., & Kuyper, H. (2010). Students' daily emotions in the classroom: Intra-individual variability and appraisal correlates. *British Journal of Educational Psychology, 80*, 583–597.

Becker, E. S., Goetz, T., Morger, V., & Ranellucci, J. (2014). The importance of teachers' emotions and instructional behavior for their students' emotions: An experience sampling analysis. *Teaching and Teacher Education, 43*, 15–26. doi: 10.1016/j.tate.2014.05.002

Bieg, M., Goetz, T., & Hubbard, K. (2013). Can I master it and does it matter? An intraindividual analysis on control-value antecedents of trait and state academic emotions. *Learning and Individual Differences, 28*, 102–108. doi: 10.1016/j.lindif.2013.09.006

Bieg, M., Goetz, T., & Lipnevich, A. A. (2014). What students think they feel differs from what they really feel – Academic self-concept moderates the discrepancy between students' trait and state emotional self-reports. *PLoS ONE, 9*, e92563. doi: 10.1371/journal.pone.0092563

Biesanz, J. C. (2012). Autoregressive longitudinal models. In R. H. Hoyle (Ed.), *Structural Equation Modeling* (pp. 459–471). New York, NY: The Guilford Press.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, 579–616. doi: 10.1146/annurev.psych.54.101601.145030

Carr, H. A., & Kingsbury, F. A. (1938). The concept of traits. *Psychological Review, 45*, 497–524. doi: 10.1037/h0063048

Cattell, R. B. (1946). *Description and measurement of personality*. Oxford, UK: World Book Company.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R: Professional Manual: Revised NEO PI-R and NEO-FFI*. Odessa, FL: Psychological Assessment Resources.

Daschmann, E. C., Goetz, T., & Stupnisky, R. H. (2011). Testing the predictors of boredom at school: Development and validation of the precursors to boredom scales. *British Journal of Educational Psychology, 81*, 421–440.

Eid, M. (1997). Happiness and satisfaction: An application of a latent state-trait model for ordinal variables. In J. Rost & R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences* (pp. 148–154). Münster, Germany: Waxmann.

Eid, M., & Diener, E. (1999). Intraindividual variability in affect: Reliability, validity, and personality correlates. *Journal of Personality and Social Psychology, 76*, 662–676. doi: 10.1037/0022-3514.76.4.662

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169–200. doi: 10.1080/02699939208411068

Epstein, S. (1977). Traits are alive and well. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp. 83–98). Hillsdale, NJ: Erlbaum.

Farmer, R., & Sundberg, N. D. (1986). Boredom proneness – the development and correlates of a new scale. *Journal of Personality Assessment, 50*, 4–17.

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology, 74*, 967–984. doi: 10.1037/0022-3514.74.4.967

Feldman Barrett, L., & Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science, 8*, 10–14.

Frenzel, A. C., & Stephens, E. J. (2013). Emotions. In N. C. Hall & T. Goetz (Eds.), *Emotion, motivation, and self-regulation: A handbook for teachers* (pp. 1–56). Bingley, UK: Emerald.

Frijda, N. H. (2007). *The laws of emotion*. Mahwah, NJ: Erlbaum.

Geiser, C., & Lockhart, G. (2012). A comparison of four approaches to account for method effects in latent state-trait analyses. *Psychological Methods, 17*, 255–283. doi: 10.1037/a002697710.1037/a0026977

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science, 24*, 2079–2087. doi: 10.1177/0956797613486989

Goetz, T., Frenzel, A. C., Hall, N. C., Nett, U. E., Pekrun, R., & Lipnevich, A. A. (2014). Types of boredom: An experience sampling approach. *Motivation and Emotion, 38*, 401–419. doi: 10.1007/s11031-013-9385-y

Goetz, T., Frenzel, A. C., Pekrun, R., & Hall, N. C. (2006). The domain specificity of academic emotional experiences. *The Journal of Experimental Education, 75*, 5–29.

Goetz, T., Frenzel, A. C., Stoeger, H., & Hall, N. C. (2010). Antecedents of everyday positive emotions: An experience sampling analysis. *Motivation and Emotion, 34*, 49–62. doi: 10.1007/s11031-009-9152-2

Goetz, T., Lüdtke, O., Nett, U. E., Keller, M. M., & Lipnevich, A. A. (2013). Characteristics of teaching and students' emotions in the classroom: Investigating differences across domains. *Contemporary Educational Psychology, 38*, 383–394. doi: 10.1016/j.cedpsych.2013.08.001

Goetz, T., Nett, U. E., Martiny, S. E., Hall, N. C., Pekrun, R., Dettmers, S., & Trautwein, U. (2012). Students' emotions during homework: Structures, self-concept antecedents, and achievement outcomes. *Learning and Individual Differences, 22*, 225–234. doi: 10.1016/j.lindif.2011.04.00

Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*, 289–308.

Hank, P. (2015). Beyond an informal everyday concept of self-esteem: A latent state-trait model. *Journal of Individual Differences, 36*, 237–246. doi: 10.1027/1614-0001/a000181

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA: Sage.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55.

Izard, C. E., Libero, D. Z., Putnam, P., & Haynes, O. M. (1993). Stability of emotion experiences and their relations to traits of personality. *Journal of Personality and Social Psychology, 64*, 847–860. doi: 10.1037/0022-3514.64.5.847

Keller, M. M., Chang, M.-L., Becker, E. S., Goetz, T., & Frenzel, A. C. (2014). Teachers' emotional experiences and exhaustion as predictors of emotional labor in the classroom: An experience sampling study. *Frontiers in Psychology, 5*, 1442.

Kenny, D. A., & Zautra, A. (2001). Trait-state models for longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 243–263) Washington, DC: American Psychological Association. doi: 10.1037/10409-008

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion, 5*, 345–379. doi: 10.1007/BF00992553

Merz, E. L., & Roesch, S. C. (2011). Modeling trait and state variation using multilevel factor analysis with PANAS daily diary data. *Journal of Research in Personality, 45*, 2–9. doi: 10.1016/j.jrp.2010.11.003

Mischel, W. (1969). Continuity and change in personality. *The American Psychologist, 24*, 1012–1018. doi: 10.1037/h0028886

Muthén, L. K., & Muthén, B. O. (1998–2013). *Mplus user's guide*, (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nesselroade, J. R. (1988). Some implications of the trait-state distinction for the study of development over the life span: The case of personality. In P. B. Baltes, D. L. Featherman, & R. M. Lerner (Eds.), *Life-span development and behavior* (Vol. 8, pp. 163–189). Hillsdale, NJ: Erlbaum.

Nett, U. E., Goetz, T., & Hall, N. C. (2011). Coping with boredom in school: An experience sampling perspective. *Contemporary Educational Psychology, 36*, 49–59. doi: 10.1016/j.cedpsych.2010.10.003

Parkinson, B. (1985). Emotional effects of false automatic feedback. *Psychological Bulletin, 98*, 471–494. doi: 10.1037/0033-2909.98.3.471

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36–48. doi: 10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002a). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*, 91–105.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002b). Positive emotions in education. In E. Frydenberg (Ed.), *Beyond coping: Meeting goals, visions, and challenges* (pp. 149–174). Oxford, UK: Elsevier.

Pekrun, R. & Linnenbrink-Garcia, L. (Eds.). (2014). *International handbook of emotions in education*. New York, NY: Routledge.

Polivy, J. (1981). On the induction of emotion in the laboratory: Discrete moods or multiple affect states? *Journal of Personality and Social Psychology, 41*, 803–817. doi: 10.1037/0022-3514.41.4.803

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*, 934–960. doi: 10.1037/0033-2909.128.6.934

Röcke, C., & Brose, A. (2013). Intraindividual variability and stability of affect and well-being: Short-term and long-term change and stabilization processes. *The Journal of Gerontopsychology and Geriatric Psychiatry, 26*, 185–199. doi: 10.1024/1662-9647/a000094

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin, 125*, 3–30. doi: 10.1037/0033-2909.125.1.3

Schimmack, U. (2003). Affect measurement in experience sampling research. *Journal of Happiness Studies, 4*, 79–106. doi: 10.1023/a:1023661322862

Spielberger C. (Ed.). (1972). *Anxiety as an emotional state*. (Vol. 1). New York, NY: Academic Press.

Spielberger, C., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Steyer, R., Ferring, D., & Schmitt, M. J. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment, 8*, 79–98.

Steyer, R., Mayer, A., Geiser, C., & Cole, D. A. (2015). A theory of states and traits – Revised. *Annual Review of Clinical Psychology, 11*, 71–98. doi: 10.1146/annurev-clinpsy-032813-153719

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality, 13*, 389–408. doi: 10.1002/(sici)1099-0984(199909/10)13:5<389::aid-per361>3.0.co;2-a

Steyer, R., Schwenkmezger, P., & Auer, A. (1990). The emotional and cognitive components of trait anxiety: A latent state-trait model. *Personality and Individual Differences, 11*, 125–134. doi: 10.1016/0191-8869(90)90004-b

Tsai, Y.-M., Kunter, M., Lüdtke, O., Trautwein, U., & Ryan, R. M. (2008). What makes lessons interesting? The role of situational and individual factors in three school subjects. *Journal of Educational Psychology, 100*, 460–472. doi: 10.1037/0022-0663.100.2.460

Vansteelandt, K., Van Mechelen, I., & Nezlek, J. B. (2005). The co-occurrence of emotions in daily life: A multilevel approach. *Journal of Research in Personality, 39*, 325–335. doi: 10.1016/j.jrp.2004.05.006

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Yasuda, T., Lawrenz, C., van Whitlock, R., Lubin, B., & Lei, P. W. (2004). Assessment of intraindividual variability in positive and negative affect using latent state-trait model analyses. *Educational and Psychological Measurement, 64*, 514–530. doi: 10.1177/0013164403258445

Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future directions. In P. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 165–183). Burlington, MA: Elsevier Academic Press.

Zelenski, J. M., & Larsen, R. J. (2000). The distribution of basic emotions in everyday life: A state and trait perspective from experience sampling data. *Journal of Research in Personality, 34*, 178–197. doi: 10.1006/jrpe.1999.2275

Ziegler, M., Ehrlenspiel, F., & Brand, R. (2009). Latent state-trait theory: An application in sport psychology. *Psychology of Sport and Exercise, 10*, 344–349. doi: 10.1016/j.psychsport.2008.12.004

Zuckerman, M. (1976). General and situation-specific traits and states: New approaches to assessment of anxiety and other constructs. In M. Zuckerman & C. D. Spielberger (Eds.), *Emotions and anxiety: New concepts, methods, and applications* (pp. 133–174). Oxford, UK: Erlbaum.

Ulrike E. Nett

Empirical Educational Research
Augsburg University
86159 Augsburg
Germany
E-mail ulrike.nett@phil.uni-augsburg.de

Madeleine Bieg

Empirical Educational Research
University of Konstanz
Universitätsstraße 10/Fach 45
78457 Konstanz
Germany
E-mail madeleine.bieg@uni-konstanz.de

Appendix

Table A1. Original item wordings and English translation for the five discrete emotions assessed with two items each

Emotion		Original German wording	English translation
Enjoyment	SJOY1	In dieser Stunde hat mir der Matheunterricht Spaß gemacht.	In this lesson mathematics was fun.
	SJOYY2	In dieser Stunde habe ich mich gefreut.	In this lesson I felt happy.
Pride	SPRD1	In dieser Stunde war ich stolz auf meine Mitarbeit.	In this lesson I was proud of my participation.
	SPRD2	In dieser Stunde war ich stolz auf mich.	In this lesson I was proud of myself.
Anxiety	SANX1	In dieser Stunde war ich beunruhigt.	In this lesson I felt uneasy.
	SANX2	In dieser Stunde hatte ich Angst.	In this lesson I felt anxious.
Anger	SANG1	In dieser Stunde war ich vor Ärger ganz unruhig.	In this lesson I felt restless due to anger.
	SANG2	In dieser Stunde habe ich mich geärgert.	In this lesson I felt angry.
Boredom	SBOR1	In dieser Stunde fand ich den Matheunterricht langweilig.	In this lesson mathematics was boring.
	SBOR2	In dieser Stunde habe ich mich gelangweilt.	In this lesson I felt bored.

Notes. Items were rated on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree.* Items were adapted from the Achievement Emotions Questionnaire (AEQ; Pekrun et al., 2011).

Table A2. Item-specific reliabilities, consistencies, and specificities

	Variance (ε_{it})	Reliability	Consistency	Specificity	% Consistency	% Specificity
Enjoyment						
SJOY1_01	0.38	0.73	0.38	0.35	51.62	48.38
SJOY2_01	0.48	0.66	0.31	0.35	47.22	52.78
SJOY1_02	0.35	0.76	0.36	0.39	48.08	51.92
SJOY2_02	0.51	0.67	0.29	0.38	43.72	56.28
SJOY1_03	0.28	0.78	0.41	0.37	52.96	47.04
SJOY2_03	0.48	0.65	0.32	0.34	48.57	51.43
SJOY1_04	0.33	0.76	0.38	0.38	50.38	49.62
SJOY2_04	0.46	0.67	0.31	0.36	45.99	54.01
SJOY1_05	0.27	0.79	0.42	0.36	53.61	46.39
SJOY2_05	0.42	0.68	0.33	0.35	49.22	50.78
SJOY1_06	0.36	0.77	0.35	0.42	45.00	55.00
SJOY2_06	0.41	0.73	0.30	0.43	40.70	59.30
Mean	0.39	0.72	0.35	0.37		
Medium	0.39	0.73	0.34	0.37		
Min	0.27	0.65	0.29	0.34		
Max	0.51	0.79	0.42	0.43		
Pride						
SPRD1_01	0.35	0.74	0.31	0.43	42.00	58.00
SPRD2_01	0.41	0.72	0.33	0.39	45.79	54.21
SPRD1_02	0.37	0.72	0.32	0.40	44.44	55.56
SPRD2_02	0.43	0.70	0.34	0.36	48.28	51.72
SPRD1_03	0.32	0.73	0.35	0.38	47.73	52.27
SPRD2_03	0.40	0.70	0.36	0.34	51.58	48.42
SPRD1_04	0.36	0.71	0.34	0.37	48.11	51.89
SPRD2_04	0.42	0.69	0.36	0.33	51.96	48.04
SPRD1_05	0.31	0.74	0.36	0.38	48.28	51.72
SPRD2_05	0.39	0.71	0.37	0.34	52.13	47.87
SPRD1_06	0.37	0.75	0.29	0.45	39.14	60.86
SPRD2_06	0.35	0.76	0.33	0.44	42.87	57.13
Mean	0.37	0.72	0.34	0.38		
Medium	0.37	0.72	0.34	0.38		
Minimum	0.31	0.69	0.29	0.33		
Maximum	0.43	0.76	0.37	0.45		
Anxiety						
SANX1_01	0.55	0.54	0.29	0.25	53.13	46.87

(Continued on next page)

Table A2. (Continued)

	Variance (ε_{it})	Reliability	Consistency	Specificity	% Consistency	% Specificity
SANX2_01	0.17	0.75	0.33	0.43	43.16	56.84
SANX1_02	0.53	0.52	0.31	0.22	58.45	41.55
SANX2_02	0.19	0.71	0.34	0.37	48.50	51.50
SANX1_03	0.43	0.59	0.33	0.26	55.21	44.79
SANX2_03	0.18	0.73	0.33	0.40	45.22	54.78
SANX1_04	0.43	0.54	0.36	0.17	68.07	31.93
SANX2_04	0.27	0.59	0.35	0.24	58.81	41.19
SANX1_05	0.43	0.56	0.35	0.21	62.66	37.34
SANX2_05	0.19	0.70	0.37	0.33	52.91	47.09
SANX1_06	0.44	0.61	0.30	0.31	48.92	51.08
SANX2_06	0.20	0.75	0.29	0.46	39.07	60.93
Mean	0.33	0.63	0.33	0.30		
Medium	0.35	0.60	0.33	0.29		
Minimum	0.17	0.52	0.29	0.17		
Maximum	0.55	0.75	0.37	0.46		
Anger						
SANG1_01	0.13	0.82	0.32	0.50	39.18	60.82
SANG2_01	0.56	0.56	0.26	0.29	47.30	52.70
SANG1_02	0.17	0.79	0.30	0.49	37.88	62.12
SANG2_02	0.52	0.58	0.27	0.32	45.93	54.07
SANG1_03	0.14	0.77	0.38	0.39	49.79	50.21
SANG2_03	0.48	0.55	0.32	0.23	58.01	41.99
SANG1_04	0.18	0.75	0.34	0.41	45.18	54.82
SANG2_04	0.53	0.54	0.29	0.25	53.45	46.55
SANG1_05	0.17	0.75	0.35	0.40	47.05	52.95
SANG2_05	0.44	0.58	0.32	0.26	55.32	44.68
SANG1_06	0.18	0.77	0.30	0.48	38.18	61.82
SANG2_06	0.49	0.60	0.28	0.32	46.25	53.75
Mean	0.33	0.67	0.31	0.36		
Medium	0.31	0.67	0.31	0.35		
Minimum	0.13	0.54	0.26	0.23		
Maximum	0.56	0.82	0.38	0.50		
Boredom						
SBOR1_01	0.34	0.76	0.36	0.40	47.23	52.77
SBOR2_01	0.25	0.81	0.38	0.43	46.61	53.39
SBOR1_02	0.24	0.83	0.38	0.45	45.50	54.50
SBOR2_02	0.24	0.83	0.37	0.46	44.88	55.12
SBOR1_03	0.35	0.75	0.38	0.37	50.73	49.27
SBOR2_03	0.27	0.79	0.40	0.39	50.10	49.90
SBOR1_04	0.33	0.76	0.38	0.38	50.00	50.00
SBOR2_04	0.30	0.78	0.38	0.39	49.37	50.63
SBOR1_05	0.25	0.80	0.42	0.38	52.05	47.95
SBOR2_05	0.25	0.80	0.41	0.39	51.42	48.58
SBOR1_06	0.26	0.84	0.33	0.50	39.80	60.20
SBOR2_06	0.28	0.82	0.32	0.50	39.20	60.80
Mean	0.28	0.80	0.38	0.42		
Medium	0.26	0.80	0.38	0.40		
Minimum	0.24	0.75	0.32	0.37		
Maximum	0.35	0.84	0.42	0.50		

Notes. Reliability, Consistency, and Specificity were calculated according to Geiser and Lockhart (2012). As all factor loadings in the models were fixed to equal 1, the formulas can be simplified to the following:

$$\text{Consistency}(E_{it}) = \frac{\text{Variance}(T_i)}{\text{Variance}(E_{it})}$$

$$\text{Specificity}(E_{it}) = \frac{\text{Variance}(S_t)}{\text{Variance}(E_{it})}$$

$$\text{Reliability}(E_{it}) = 1 - \frac{\text{Variance}(\varepsilon_{it})}{\text{Variance}(E_{it})} = \text{Consistency}(E_{it}) + \text{Specificity}(E_{it})$$

The values of % Consistency and % Specificity contain the percentages of the consistency and specificity scores of the explained variance which equals the reliability score. Thus, % Consistency and % Specificity add up to 100%.

Table A3. Fit indices of single latent trait multilatent state models

	Chi-square; *df* (*p*)	CFI/TLI	RMSEA	SRMR
Enjoyment	255.18, 58, (< .001)	0.93/0.92	0.06	0.05
Pride	138.05, 58, (< .001)	0.97/0.97	0.04	0.05
Anxiety	238.46, 58, (< .001)	0.88/0.87	0.06	0.06
Anger	129.58, 58, (< .001)	0.96/0.95	0.03	0.05
Boredom	132.95, 58, (< .001)	0.98/0.98	0.04	0.05

Notes. Loadings for each specific item were fixed to be equal across all measurement points for the latent trait. Loadings for latent state residuals were fixed to equal 1.