

The adaptation and validation of the Achievement Emotions Questionnaire for Physical Education-Elementary School

Mehmet Ceylan ^{1*}, Kevser Ceylan ¹, Oğuzhan Yoncalık ¹

¹Kırıkkale University, Faculty of Sport Sciences, Department of Physical Education and Sport, Kırıkkale, Türkiye

ARTICLE HISTORY

Received: June 30, 2025

Accepted: Feb. 3, 2026

KEYWORDS

Achievement emotions,
Physical education,
Elementary school,
Control-value theory,
Visual scale.

Abstract: Achievement emotions such as enjoyment, anxiety, and boredom significantly influence students' motivation, engagement, and performance in physical education (PE). Despite growing interest in academic emotions, age-appropriate tools for younger students remain limited. This study aimed to adapt and validate a brief, visually supported scale to assess enjoyment, anxiety, and boredom in PE. The goal was to develop a developmentally suitable instrument grounded in structural, linguistic, and cultural equivalence. Participants were 772 third- and fourth-grade students (388 girls) from six public elementary schools, recruited through convenience sampling. The adaptation process involved forward-backward translation, expert evaluations to ensure semantic, idiomatic, experiential, and conceptual equivalence, a pre-test to verify item clarity, and a pilot study to refine the visual layout. Psychometric properties were examined using Confirmatory Factor Analysis and Exploratory Structural Equation Modeling (ESEM). Results supported a clear three-factor structure with strong model fit indices. ESEM yielded high factor loadings and minimal cross-loadings, confirming distinct constructs. Internal consistency was satisfactory, with Cronbach's alpha and McDonald's omega values above .70. Construct validity was supported by acceptable AVE and CR values, and discriminant validity was confirmed. Multi-group CFA demonstrated measurement invariance across gender (male-female) and across grade level (3rd vs. 4th), indicating stable functioning across these groups. Overall, the findings show that the adapted scale is a reliable, valid, and practical instrument for assessing achievement emotions in elementary PE, offering a brief and visually accessible option for evaluating young students' emotional experiences.

INTRODUCTION

The educational process is a multidimensional structure that involves not only cognitive gains but also emotional experiences. In this context, physical education classes are considered especially important for emotional development rather than cognitive outcomes (Simonton & Garn, 2019). In fact, many experts today argue that physical education should be approached as an emotionally driven subject (Espoz-Lazo *et al.*, 2020; Wright *et al.*, 2020). Like many other subjects, physical education also takes place in an achievement-oriented environment. The emotions that students experience before, during, or after class in connection with this achievement setting are referred to as achievement emotions (Simonton & Garn, 2019).

*Corresponding author: Mehmet CEYLAN ✉ mhmtcyln183@gmail.com 📍 Kırıkkale University, Faculty of Sport Sciences, Department of Physical Education and Sport, Kırıkkale, Türkiye

The copyright of the published article belongs to its author under CC BY 4.0 license. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>

e-ISSN: 2148-7456

Achievement emotions are emotional responses that can be either positive or negative. They are linked to students' academic performance and are shaped by their perceptions and experiences of success. These emotions influence both learning and motivation (Pekrun, 2006, 2024) and, therefore, play a critical role in academic achievement (Camacho-Morles *et al.*, 2021). In addition to their impact on performance, it is now well understood that these emotions also affect important aspects of life, such as well-being, health, and career decisions (Pekrun *et al.*, 2023). For this reason, understanding how achievement emotions are formed and how they affect students' learning processes is essential. To understand the development of these emotions, we can look to Pekrun's (2006) Control-Value Theory. This theory suggests that achievement emotions are shaped by two main factors: control and value. Control refers to a student's belief in their ability to influence their actions and outcomes in the learning process. Value refers to the importance and meaning the student attaches to the task. In physical education, factors such as the teacher, classroom environment, perceived physical competence, and peer support can directly shape students' emotional experiences (Simonton *et al.*, 2024).

Although school is inherently an emotional environment (Pekrun, 2024), research on emotions in educational settings, especially in physical education, remains limited (Fierro-Suero *et al.*, 2020; Simonton & Garn, 2019). In physical education, previous work (e.g., Fierro-Suero *et al.*, 2020; Simonton & Garn, 2019) has already noted that achievement emotions are underexplored in general PE contexts, and this gap appears to be even more pronounced at the elementary level, where domain-specific studies are relatively scarce. In physical education classes, achievement emotions play a key role in shaping students' emotional connection to the subject, their motivation, and their tendency to engage in physical activity over the long term (Fierro-Suero *et al.*, 2020; Garn *et al.*, 2017; Simonton & Garn, 2019). Positive achievement emotions strengthen students' participation in class and enhance their intrinsic motivation toward physical activity (Ceylan & Karlı, 2024; Simonton & Garn, 2020). In contrast, negative achievement emotions may increase students' anxiety toward the subject and, over time, lead to avoidance behaviors (Bieleke *et al.*, 2023; Cañabate *et al.*, 2018). It is important to systematically investigate achievement emotions at the elementary level, especially in physical education classes where the foundations of education are laid, and interest in lifelong physical activity is cultivated.

Although the importance of achievement emotions has been emphasized and the elementary school period is considered a critical stage for the development of these emotions, there is currently no valid and reliable measurement tool available for assessing them in physical education classes, which contribute significantly to both physical and psychological health. A review of the achievement emotions literature shows that various measurement instruments have been developed for use in educational settings (e.g., Bieleke *et al.*, 2021; Lichtenfeld *et al.*, 2012; Pekrun *et al.*, 2011). There are also tools designed to measure achievement emotions specifically in the context of physical education (e.g., Ceylan & Karlı, 2024; Fierro-Suero *et al.*, 2020; Simonton *et al.*, 2023). However, none of these scales target the elementary level in physical education. This is a critical gap, as the cognitive and developmental characteristics of elementary school students differ significantly from those of older students. Scales designed for adolescents (e.g., Ceylan & Karlı, 2024; Fierro-Suero *et al.*, 2020) often rely on text-based items that require a level of abstract self-reflection and reading comprehension that may not be appropriate for younger children. Accordingly, in the present adaptation, we did not only use a visually supported response format; we also reformulated the item wording to reflect the movement- and play-based nature of elementary physical education and to match children's everyday language. In contrast, elementary students often express their emotional experiences more easily and accurately when provided with concrete, visually supported formats (Lichtenfeld *et al.*, 2012). In one notable example, Lichtenfeld *et al.* (2012) developed the Achievement Emotions Questionnaire-Elementary School (AEQ-ES) to assess achievement emotions in mathematics among elementary school students. Although the AEQ-ES was

designed in a traditional Likert-type format, it was adapted to the cognitive level of younger children by incorporating facial expressions to represent emotions. This visual format was intended to help students express their emotional experiences more easily and accurately (Lichtenfeld *et al.*, 2012).

From a test-adaptation perspective, structural equivalence across cultures and languages is essential (Hambleton & Patsula, 1999). In the present study, we assumed that the latent achievement emotions of enjoyment, anxiety, and boredom have clear counterparts in both the original AEQ-ES context and the Turkish PE context, and that cross-cultural comparison is meaningful because these emotions are defined within the same control-value theoretical framework. Previous research has consistently described these emotions as distinct but related dimensions, typically modeled as a three-factor structure at the elementary level (e.g., Lichtenfeld *et al.*, 2012 in mathematics; Ceylan & Karlı, 2024; Fierro-Suero *et al.*, 2020; Simonton *et al.*, 2023 in PE). By retaining this three-emotion structure and testing it with CFA, ESEM, and measurement invariance analyses, we sought to provide empirical evidence that structural equivalence has been adequately achieved and verified in the adapted AEQ-PE-ES.

The Current Study

This study aimed to develop and validate the Achievement Emotions Questionnaire for Physical Education-Elementary School (AEQ-PE-ES). Rather than constructing a completely new instrument, we intentionally adapted the Achievement Emotions Questionnaire-Elementary School (AEQ-ES; Lichtenfeld *et al.*, 2012) to the physical education domain. The AEQ-ES is one of the foundational measures for assessing enjoyment, anxiety, and boredom in elementary students, is firmly grounded in control-value theory, and has demonstrated strong reliability, validity, and cross-cultural measurement invariance in German and American samples. In line with test-adaptation guidelines (Hambleton & Patsula, 1999), adaptation is recommended when the underlying construct is conceptually equivalent across populations and when comparability with existing research is desired. In our case, we preserved the original three-emotion structure, age-appropriate visual response format, and control-value-based conceptualization of the AEQ-ES, while adapting item content from mathematics to physical education. This approach is consistent with previous PE achievement emotion scales that have adapted Pekrun's achievement emotion measures to specific subjects (e.g., from mathematics to PE; Fierro-Suero *et al.*, 2020) and positions the AEQ-PE-ES as a domain-specific extension of a theoretically and empirically established instrument rather than an entirely new, isolated measure. Furthermore, evidence regarding the cultural suitability of this structure is supported by recent studies in Turkey (e.g., Ceylan & Karlı, 2024), which have successfully validated the applicability of Control-Value Theory constructs within the Turkish educational context. These studies suggest that the three-factor emotional structure is cross-culturally stable and relevant for Turkish students.

The AEQ-PE-ES is grounded in Pekrun's (2006) domain-general Control-Value Theory, which has been established as a valid framework for understanding achievement emotions in physical education (e.g., Simonton & Garn, 2019). Although other scales for physical education exist (e.g., Fierro-Suero *et al.*, 2020), they typically rely on abstract text-only response anchors and target older students, making them cognitively unsuitable for elementary students who benefit more from visually supported response formats, as demonstrated by Lichtenfeld *et al.*, (2012). Therefore, the present study drew on the pioneering work of Lichtenfeld *et al.*, (2012), who successfully measured enjoyment, anxiety, and boredom in an elementary school mathematics context using an age-appropriate, visually supported format, and extended this approach to the PE domain. In order to examine the structural dimensions of achievement emotions more comprehensively, the study employed both a traditional method, Confirmatory Factor Analysis (CFA), and a more advanced approach, Exploratory Structural Equation Modeling (ESEM), which has gained prominence in recent years for evaluating constructs in psychological

research. Compared to conventional CFA, the ESEM technique offers a more flexible modeling framework (van Zyl & ten Klooster, 2022). Strictly constraining cross-loadings to zero, as in CFA, may not align with the developmental reality of elementary school children, whose emotional boundaries are often less differentiated and more intertwined than those of adults (Lichtenfeld *et al.*, 2012). Furthermore, from a cultural adaptation perspective, ESEM allows for the detection of potential semantic overlaps that may arise when translating emotional concepts into the Turkish context, preventing artificial model misfit due to minor cross-cultural variations (van Zyl & ten Klooster, 2022). The findings of this study are expected to provide a deeper understanding of the structural characteristics of achievement emotions in physical education (see Figure 1) and contribute to our knowledge of how these emotions develop at an early age.

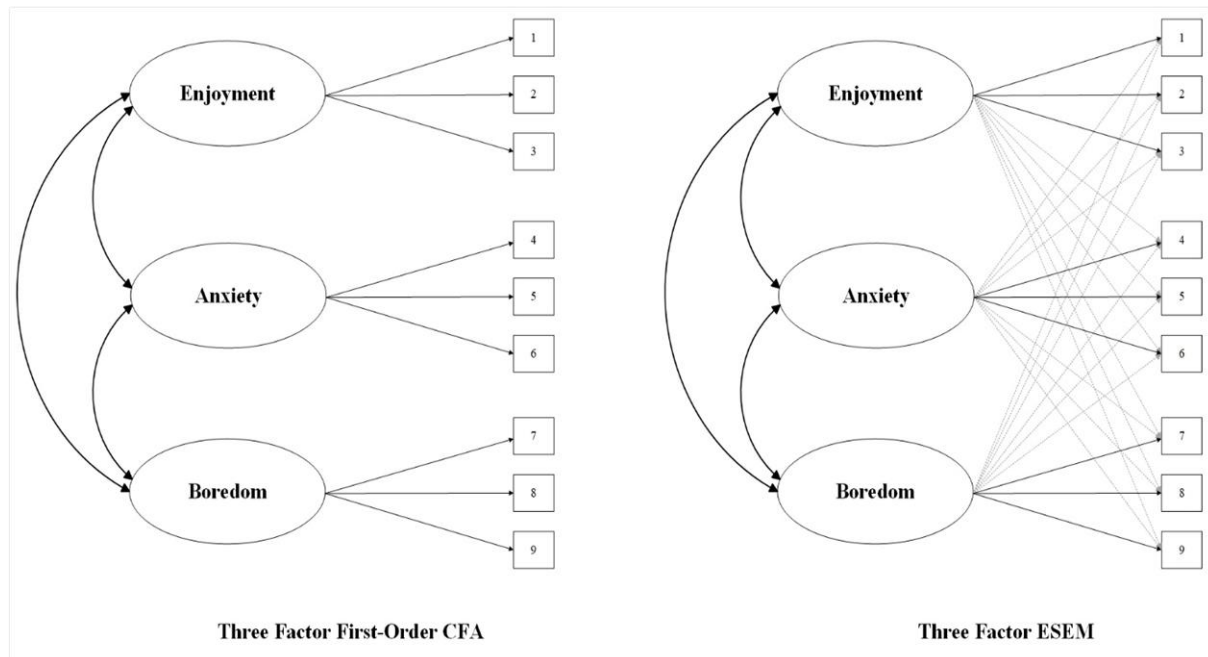


Figure 1. Graphical representation of the tested models.

METHOD

Participants and Procedure

This study employed a cross-sectional survey research design to adapt and validate the scale (Fraenkel *et al.*, 2012). This study involved third and fourth-grade students attending six different public elementary schools affiliated with the Ministry of National Education in the province of Kırıkkale, Türkiye. The sample consisted of a total of 772 students (388 female) from these schools. Of the participants, 380 were in the third grade, and 392 were in the fourth grade. The target population was deliberately chosen to mirror the original AEQ-ES study, which also focused on elementary school students in the early grades (Lichtenfeld *et al.*, 2012). Although the exact sample sizes and national contexts differ, both studies draw on large groups of children from public schools in general education settings, making the present sample reasonably comparable to the original instrument's target group. Any differences in sample size and context were considered a strength rather than a limitation, as they allowed us to test the robustness of the AEQ-ES structure in a new cultural and domain-specific context. The students were between the ages of 9 and 11 ($M = 9.61$, $SD = .63$). Although various rules of thumb exist, Pituch & Stevens (2015) recommend a minimum of 20 participants per variable (item) for the sample size. Based on this criterion, our final sample of 772 participants substantially meets this requirement, providing sufficient statistical power for the CFA and ESEM analyses. In addition, 110 students were administered the data twice for test-retest analysis.

Participation in the study was entirely voluntary, and only students whose parents provided written consent were included in the research. During data collection, the purpose of the study was explained to the students, and they were informed that the data would be used solely for scientific purposes. It was clearly stated that there were no right or wrong answers on the questionnaire and that students were simply expected to express their personal feelings and thoughts. Considering the age range of the sample, each item on the scale was read aloud in the classroom by the lead researcher to ensure that every student fully understood the questions. Completing the scale took approximately 15 minutes and was done under teacher supervision.

Measurement

In this study, the AEQ-ES scale developed by Lichtenfeld *et al.* (2012) was adapted into Turkish and tailored specifically to the context of elementary school physical education classes. The AEQ-ES was originally developed within the framework of the Control-Value Theory and aims to assess three core achievement emotions experienced by students in academic settings: enjoyment, anxiety, and boredom. This specific focus on three dimensions, rather than a larger set, is theoretically deliberate. According to Pekrun's (2006) Control-Value Theory, enjoyment, anxiety, and boredom are the three most central and prevalent activity-related emotions (i.e., emotions experienced during the learning activity). In contrast, other emotions often included in broader scales (e.g., pride, shame, hopelessness) are classified as outcome-related emotions (i.e., felt in response to a success or failure). Given that the present study intentionally excluded outcome contexts (such as tests and homework, as mentioned later in this section) to focus exclusively on the in-class experience of PE, limiting the scale to these three core activity-related emotions was a necessary and theoretically consistent decision to maintain construct precision. While the original version of the scale measures these emotions in the contexts of class, homework, and tests, the present study focused solely on students' overall emotional experiences related to physical education. This choice was based on the structural characteristics of physical education classes and the developmental level of elementary school students, which led to the exclusion of testing and homework contexts from the adapted version.

The scale items were reformulated to reflect the physically active and play-based social nature of physical education classes. The adapted scale consists of a total of nine items, with three items for each emotion. The number of items was kept limited in order to ensure that the scale would be appropriate for the attention span of the target age group and that it would be easy to understand and administer (DeVellis & Thorpe, 2021). Each item was rated using a five-point Likert scale (not at all, very little, somewhat, quite a bit, very much). To accommodate the age characteristics of the students, the response options were visually supported with facial expressions (emojis). This method was also used in the original version of the AEQ-ES, allowing children to express their emotions more concretely by associating them with facial expressions that represent increasing emotional intensity (Lichtenfeld *et al.*, 2012). In addition, different character faces were designed for boys and girls to help students better identify with the facial expressions. All faces were structured consistently to convey the same emotional meaning and were redesigned in a clear and simple visual format specifically for this study (see [Appendix 1](#)). The translation was conducted by experts proficient in both languages and cultures and knowledgeable about the construct (Sousa & Rojjanasrirat, 2011); thus, specific training was not required. The resulting scale was reviewed by three academic experts. Two of these experts are senior researchers with over 15 years of experience, specifically specializing in achievement emotions and sport psychology. The third expert holds a PhD in physical education with five years of academic experience and possesses advanced English proficiency, ensuring linguistic accuracy. They evaluated the scale in terms of (a) semantic, idiomatic, experiential, and conceptual equivalence, (b) how well each item represented the structure of achievement emotions, (c) the clarity of the expressions, and (d) whether any items appeared to be repetitive in meaning. The scale was also piloted with 24 elementary school students (13

girls), who were asked whether they understood the meaning of each item. Based on the results of the pilot study, it was concluded that the scale was suitable for use in the main research.

Data Analysis

Initially, data were collected from 780 participants. After the research data were transferred to a digital format, checks for missing values were performed. Responses with a completion rate below 95% were handled using listwise deletion. A total of 8 participants were removed based on this criterion, leaving the final sample of 772 included in the analyses. There were no missing data points in the records of the remaining 772 individuals; thus, the final dataset was complete. Descriptive statistics, including means, standard deviations, skewness, and kurtosis coefficients, were calculated using SPSS 27. The assumption of normal distribution for each subscale was evaluated based on whether the skewness and kurtosis values remained within the range of ± 1 . For distributions that exceeded this threshold, the Robust Maximum Likelihood (MLR) estimator was applied (van Zyl & ten Klooster, 2022).

To test the construct validity of the scale, both CFA and ESEM were conducted. Whereas CFA was used to test the hypothesized simple three-factor structure, ESEM was employed to provide a more flexible evaluation of the measurement model by allowing items to cross-load on non-target factors. This flexibility is critical because strictly constraining these naturally occurring conceptual overlaps to zero, as in CFA, often results in inflated factor correlations and biased structural estimates (Marsh *et al.*, 2014; van Zyl & ten Klooster, 2022). This approach helps to obtain more realistic factor correlations and offers stronger evidence for convergent and discriminant validity of the three achievement emotion dimensions (van Zyl & ten Klooster, 2022). These two analyses were performed using the Mplus 8.3 software (Muthén & Muthén, 2017). In the CFA, items were allowed to load only onto their predefined factors, and cross-loadings were not permitted. In the ESEM analysis, items were allowed to load freely onto their designated factors, while loadings on non-target factors were constrained to remain as close to zero as possible through the use of target rotation. This approach allows for more flexible modeling of psychological constructs such as achievement emotions (Morin *et al.*, 2016). To assess model fit, we reported a comprehensive set of indices, including χ^2 , TLI (Tucker-Lewis Index), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). CFI values equal to or greater than .90, and RMSEA and SRMR values equal to or less than .08, were considered acceptable thresholds (Hu & Bentler, 1999). In addition, AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), and aBIC (Adjusted BIC) values were examined to compare the models. Lower values on these indices indicate better model fit with lower complexity (Lin *et al.*, 2017).

To assess the structural equivalence of the scale across gender and grade level, a measurement invariance analysis was conducted. This analysis was performed at the configural, metric, scalar, and strict levels. A change in CFI (Δ CFI) of less than .01 between models was taken as the criterion for establishing measurement invariance (Byrne, 2016). To evaluate internal consistency, Cronbach's Alpha and McDonald's Omega (ω) coefficients were calculated. Values above .70 for both coefficients were considered indicative of acceptable reliability (Reise *et al.*, 2013). To assess construct validity, Average Variance Extracted (AVE) and Composite Reliability (CR) values were computed. The criteria for construct validity were set as $AVE \geq .50$ and $CR \geq .70$. In cases where the AVE value was below the threshold, but the CR was sufficient, construct validity was still considered acceptable (Fornell & Larcker, 1981). For discriminant validity, the Maximum Shared Squared Variance (MSV) and Average Shared Squared Variance (ASV) were used. Discriminant validity was considered to be established when $AVE > MSV$, $AVE > ASV$, and $\sqrt{AVE} >$ inter-factor correlations (Fornell & Larcker, 1981). Finally, in order to evaluate the consistency of the scale over time, the test-retest method was applied at two-week intervals, and the coefficients obtained enabled us to evaluate whether the scale was reliable over time (Streiner *et al.*, 2024).

RESULTS

Table 1 presents the correlation values and descriptive statistics for the AEQ-PE-ES scale. Significant relationships were observed among all subscales. The strongest negative correlation was found between enjoyment and boredom ($r = -.496$), while the strongest positive correlation was between anxiety and boredom ($r = .440$). Additionally, the internal consistency coefficients for the subscales ranged from .704 to .798.

Table 1. Descriptive statistics and correlations of the AEQ-PE-ES.

	Mean	SD	Skewness	Kurtosis	Enjoyment	Anxiety	Boredom
Enjoyment	4.523	.560	-1.664	3.490	(.704)		
Anxiety	1.871	.850	1.139	.882	-.340	(.792)	
Boredom	1.559	.681	1.651	3.003	-.496	.440	(.707)
ω	—	—	—	—	.706	.798	.712

Note. Values in parentheses indicate Cronbach's Alpha coefficients.

As shown in **Table 2**, both the CFA and ESEM models demonstrated excellent fit indices. The model fit values obtained from the CFA analysis were excellent ($\chi^2_{(24)} = 33.013$, TLI = .990, CFI = .993, RMSEA = .022, SRMR = .026). Similarly, the ESEM model also showed an excellent fit ($\chi^2_{(12)} = 11.339$, TLI = 1.000, CFI = 1.000, RMSEA = .000, SRMR = .010). When AIC, BIC, and aBIC values were examined, both models yielded very similar results. Despite minimal differences, both models fell within acceptable ranges and met the established criteria supporting the factorial validity of the adapted scale.

Table 2. Results of model fit comparison between CFA and ESEM.

Type of Model	χ^2	df	TLI	CFI	RMSEA	SRMR	AIC	BIC	aBIC	Meets Criteria
CFA	33.013	24	.990	.993	.022	.026	15479.60	15619.07	15523.81	Yes
ESEM	11.339	12	1.000	1.000	.000	.010	15473.25	15668.53	15535.16	Yes

Although the models demonstrated excellent fit indices, comparing model fit values alone is not sufficient to determine the most appropriate model. Morin and colleagues (2016, 2017) emphasized that, in addition to fit indices, a careful examination of each factor loading plays a critical role in understanding how well the model represents the underlying structure. Therefore, in order to evaluate the factor structures, the factor loadings and cross-loadings are presented in **Table 3**. An examination of the factor loadings for the CFA model presented in **Table 3** shows that each item loaded strongly and significantly onto its intended factor. The ESEM analysis, which also considers cross-loadings, revealed that item 6 loaded onto the boredom factor at a very low but statistically significant level. Both models produced significant and strong factor loadings overall. Importantly, the ESEM solution showed strong primary loadings and only small cross-loadings, indicating that each item primarily measured its intended emotion while still capturing the natural overlap between enjoyment, anxiety, and boredom, thereby supporting the discriminant validity of the three-factor structure.

Table 4 shows the results of the measurement invariance analysis. When the single group solutions were analyzed, the three-factor model showed a very good fit for boys ($\chi^2_{(24)} = 52.260$, TLI = .959, CFI = .973, SRMR = .040, RMSEA = .055) and for girls ($\chi^2_{(24)} = 36.305$, TLI = .981, CFI = .988, SRMR = .030, RMSEA = .036). In the multi-group CFA results, the Δ CFI values between the configural, metric, scalar, and strict models were all below .01, indicating that the scale meets the criteria for measurement invariance across gender. These results suggest that the scale reflects the same factor structure for both male and female students.

Table 3. Standardized factor loadings.

Indicators	CFA	ESEM		
	λ	Enjoyment λ	Anxiety λ	Boredom λ
Enjoyment				
1	.729	.831	.056	-.024
2	.621	.539	-.023	.003
3	.663	.603	-.045	-.030
Anxiety				
4	.775	.020	.745	-.058
5	.827	-.021	.810	-.016
6	.643	-.008	.687	.159
Boredom				
7	.758	.072	.089	.786
8	.721	-.077	.012	.739
9	.560	-.053	-.093	.435

Table 4. Measurement invariance results.

Models	χ^2 (df)	TLI	CFI	SRMR	RMSEA	Model Comparison		
							$\Delta\chi^2$ (Δdf)	ΔCFI
Single group solutions								
Male (384)	52.260 (24)	.959	.973	.040	.055	—	—	—
Female (388)	36.305 (24)	.981	.988	.030	.036	—	—	—
3 rd Grade (380)	43.139 (24)	.972	.981	.036	.046	—	—	—
4 th Grade (392)	42.135 (24)	.971	.981	.036	.044	—	—	—
Gender (male-female) invariance								
1. Configural	88.566 (48)	.970	.980	.036	.047	—	—	—
2. Metric	99.475 (54)	.970	.978	.045	.047	2 vs. 1	10.91 (6)	.002
3. Scalar	105.834 (60)	.973	.977	.046	.044	3 vs. 2	6.359 (6)	.001
4. Strict	123.508 (69)	.972	.973	.051	.045	4 vs. 3	17.674 (9)	.004
Grade (3 rd - 4 th) invariance								
1. Configural	85.274 (48)	.972	.981	.036	.045	—	—	—
2. Metric	90.752 (54)	.975	.981	.041	.042	2 vs. 1	5.478 (6)	.000
3. Scalar	101.945 (60)	.975	.979	.044	.043	3 vs. 2	11.193 (6)	.002
4. Strict	150.428 (69)	.957	.959	.056	.055	4 vs. 3	48.483 (9)	.020

In addition, measurement invariance was also examined across grade level (3rd vs. 4th grade). In the single-group CFAs, the three-factor model demonstrated good fit for 3rd-grade students ($n = 380$; $\chi^2(24) = 43.139$, TLI = .972, CFI = .981, SRMR = .036, RMSEA = .046) and for 4th-grade students ($n = 392$; $\chi^2(24) = 42.135$, TLI = .971, CFI = .981, SRMR = .036, RMSEA = .044). In the multi-group CFA results, configural, metric, and scalar invariance across grades were supported, whereas strict invariance was not fully supported ($\Delta\text{CFI} = .020$). Overall, these findings indicate that the scale measures the same latent constructs comparably across 3rd- and 4th-grade students at least up to the scalar level.

Table 5 reports the convergent, discriminant, and reliability values. The AVE values were calculated as .452 for Enjoyment, .566 for Anxiety, and .469 for Boredom. When AVE is below .50, some sources suggest that structural validity can still be considered acceptable if the CR value exceeds .70. It was observed that the CR values were sufficient for all factors. In addition, for each factor, the AVE value was greater than the corresponding MSV and ASV values, and $\sqrt{\text{AVE}}$ was greater than the inter-factor correlations. This indicates that discriminant validity was established. Finally, the scale's consistency over time was examined at the factor level ($n = 110$), revealing high test-retest reliability coefficients for all three factors: enjoyment ($r = .887$), anxiety ($r = .878$), and boredom ($r = .827$).

Table 5. Construct validity indicators.

Factors	AVE	CR	MSV	ASV	Test-retest
Enjoyment	.452 (.672)	.711	.246	.181	.887
Anxiety	.566 (.752)	.795	.194	.155	.878
Boredom	.469 (.685)	.723	.246	.220	.827

Note. The values in parentheses represent the $\sqrt{\text{AVE}}$ values.

DISCUSSION

The aim of this study was to examine the validity and reliability of the AEQ-PE-ES, a scale adapted to measure the achievement emotions experienced by elementary school students during physical education classes. This scale was adapted from the AEQ-ES developed by Lichtenfeld *et al.* (2012) and focuses on assessing three key achievement emotions- enjoyment, anxiety, and boredom- within the context of physical education. Although this study involved a Turkish adaptation, our primary aim was to extend the AEQ-ES from an elementary mathematics context to elementary physical education. Because the scale is grounded in Control-Value Theory and uses developmentally appropriate, concrete wording with visually supported response options, we expected the three-emotion structure (enjoyment, anxiety, boredom) to generalize to PE lessons in Turkish elementary schools. The findings revealed that the construct validity of the AEQ-PE-ES was supported by high model fit values obtained from both CFA and ESEM analyses. In both methods, the model fit indices met the recommended threshold levels. The internal consistency of the scale, measured by Cronbach's Alpha and McDonald's Omega, was found to be within acceptable limits (Reise *et al.*, 2013). Additionally, the AVE and CR values calculated for construct validity were at or above the thresholds proposed by Fornell and Larcker (1981). Measurement invariance analyses indicated that the scale functioned consistently across male and female groups and across grade level (3rd vs. 4th grade), showing that it can be confidently used in comparative analyses. Overall, these results demonstrate that the brief and visually supported format of the AEQ-PE-ES effectively captures the emotional experiences of elementary school students in physical education settings.

The validity and reliability findings of the AEQ-PE-ES are largely consistent with similar scale studies in the literature. When first compared with the AEQ-ES developed by Lichtenfeld *et al.* (2012), both studies demonstrated strong construct validity supported by high model fit values. While factor loadings in the study by Lichtenfeld *et al.* ranged from .61 to .93, in the present

study they were found to range from .56 to .83. This indicates that the items in both studies adequately represent their corresponding emotional dimensions (Hu & Bentler, 1998). In Lichtenfeld *et al.*'s study, significant correlations were found between enjoyment and anxiety (-.48), enjoyment and boredom (-.76), and anxiety and boredom (.57). In the current study, the correlations were observed as -.34, -.50, and .44, respectively. These similarities suggest that both versions of the scale reflect the structural relationships between emotions in a consistent manner. Importantly, this convergence between the original and Turkish versions demonstrates that the evaluation-based structure proposed by Control-Value Theory can be generalized in linguistic and educational contexts when the meanings of the elements are preserved. However, the slight variation in the magnitude of these correlations likely reflects domain-specific differences; unlike the solitary nature of mathematics tests, PE is a public performance setting where social comparison may distinctively influence the intensity of emotional interplay.

On the other hand, when the findings of this study are compared with those of the AEQ-PE developed by Fierro-Suero *et al.* (2020) for middle school students, both studies show strong construct validity supported by high model fit indices. In their study, factor loadings ranged from .47 to .83, while in the present study, they ranged from .56 to .83. This indicates that in both studies, the items sufficiently and similarly represented their respective emotional dimensions. In their study, the reported correlations were -.24 between enjoyment and anxiety, -.71 between enjoyment and boredom, and .20 between anxiety and boredom. Similarly, in our study, the corresponding values were -.34, -.50, and .44. These results suggest that the AEQ-PE and AEQ-PE-ES scales exhibit similar structural and relational patterns, and support the applicability of the control-value theory across different age groups. These results indicate that the AEQ-PE and AEQ-PE-ES scales exhibit similar structural and relational models and support the applicability of control-value theory across different age groups. Notably, these similarities emerge despite being obtained in different cultural-educational settings, consistent with the theory's assumption that control and value assessments can evoke similar feelings of success in different contexts.

The ESEM findings are consistent with results reported in other recent scale development studies, such as the Discrete Emotions in Physical Education Scale (DEPES) by Simonton *et al.* (2023) and the Achievement Emotion Questionnaire-Second Language Learning (AEQ-L2L) by Shao *et al.* (2023). For example, in the DEPES study, Simonton *et al.* presented the ESEM approach as a more flexible and explanatory model compared to the stricter assumptions of CFA, allowing for a more realistic representation of overlaps between factors (Simonton *et al.*, 2023). Similarly, in the present study, the ESEM model showed high fit indices while maintaining cross-loadings at low levels through the target rotation framework. This allowed for a more realistic estimation of factor overlaps without compromising the structural integrity of the individual emotions. In the DEPES study, inter-factor correlations were also reported at high levels, such as -.610 between enjoyment and shame, and .749 between enjoyment and pride. These findings structurally parallel the present study's results, such as -.50 between enjoyment and boredom, and -.34 between enjoyment and anxiety. Likewise, in the AEQ-L2L study by Navarro *et al.* (2024), the ESEM model was preferred, and cross-loadings were found to be low ($\lambda < .05$), while item loadings on their intended factors were high ($\lambda > .05$). These findings support the strong construct validity and internal consistency observed in the present study and demonstrate that a multidimensional emotional structure can be effectively measured using brief and targeted items.

An examination of the ESEM results revealed that item 6 of the anxiety subscale showed a statistically significant but low loading ($\lambda = .159$) on the boredom factor. This finding suggests that it may stem from the multidimensional and overlapping nature of achievement emotions. According to Pekrun's (2006) Control-Value Theory, both anxiety and boredom are associated with low perceptions of control and value. This implies that students may sometimes experience these two emotions simultaneously. Especially in elementary school children, the boundaries

between emotions may be more fluid, which can cause some items to evoke associations with more than one emotional domain (Lichtenfeld *et al.*, 2012). Furthermore, the specific content of item 6 (I am afraid of not being able to do a movement) provides a more direct explanation for this cross-loading. The fear of failing at a specific movement (low perceived competence) may lead a child to disengage from the task, perceiving the anxiety-inducing situation as overwhelming or dull. This sense of being overwhelmed is closely aligned with the concept of boredom and is, in fact, represented by another item (Item 8) within our scale's Boredom factor. Similarly, in the DEPES study by Simonton *et al.* (2023), low-level cross-loadings were also observed and were attributed to the idea that children's emotions may not yet be cognitively well differentiated. In this context, the partial loading of an anxiety item on the boredom factor can be explained by the possibility that students perceive anxiety-inducing situations as simultaneously dull or overwhelming. This finding, both theoretically and in line with previous research, reflects the intertwined structure of children's emotional experiences.

Beyond the psychometric results, these findings have significant implications for both educational practice and future research. From a practical perspective, the AEQ-PE-ES offers teachers a quick and cognitively appropriate tool to monitor the 'emotional climate' of their classes. Given the scale's visual nature, practitioners can easily identify students experiencing high anxiety or boredom without disrupting the flow of the lesson. This allows for timely pedagogical interventions such as adjusting task difficulty or modifying game rules to foster enjoyment and prevent early disengagement from physical activity. For researchers, this study fills a critical measurement gap by providing a valid instrument specifically tailored for the elementary level, a period often overlooked in PE emotion research. Theoretically, the validation of the three-factor structure primarily supports the cross-contextual robustness of the Control-Value Theory framework in PE, even though cross-cultural invariance has not been directly tested in the present study. Methodologically, the successful application of the visual format confirms that abstract emotional constructs can be reliably measured in younger populations when age-appropriate supports (emojis) are utilized, encouraging future researchers to adopt similar visual approaches for elementary cohorts.

Limitations and Future Research

This study has several limitations. First, the research was conducted only with third- and fourth-grade students from selected schools in a single province using a convenience sampling method. This limits the generalizability of the findings to the broader student population. Similar studies conducted with different age groups (such as first- and second-grade students) and in diverse socio-economic and socio-cultural contexts could provide stronger evidence for the validity of the scale. Second, the measurement relied solely on students' self-reports. In particular, the limited ability of younger children to interpret and express their emotions may be a significant factor affecting data reliability. Therefore, future research is recommended to use mixed methods supported by multiple data sources, such as teacher or parent evaluations. Third, although the AVE values for the Enjoyment and Boredom factors were below the .50 threshold, convergent validity was still deemed acceptable because the CR values were above the .70 criterion; however, this situation was noted as a minor limitation. Fourth, this study focused only on the scale's internal structure and did not provide external validity evidence (e.g., correlations with other constructs). This should be addressed in future research. Finally, the scale is limited to three emotional dimensions. Future studies may develop extended models by including additional outcome-related achievement emotions such as pride, shame, or hopelessness.

Conclusion

This study introduced the AEQ-PE-ES as a visually supported, valid, and reliable measurement tool for assessing achievement emotions specific to elementary school physical education classes. The scale is structured within the framework of the Control-Value Theory and evaluates

the emotions of enjoyment, anxiety, and boredom in a brief and comprehensible manner. The findings indicate that the scale is valid, reliable, and practical for use. Considering the challenges of measuring emotions in younger age groups, this scale provides a valuable contribution to both physical education research and practice. It is recommended that future studies explore the use of this scale in different contexts, such as cross-cultural comparisons, intervention programs, and longitudinal research.

Acknowledgments

The authors would like to express their sincere gratitude to the school administrators, the classroom teachers who supervised the data collection, the physical education teachers who facilitated access to the participating schools, and the students who participated in this study. We also thank the Ministry of National Education (Approval No: MEB.TT.2025.019856) and the Ethics Committee of Kırıkkale University (Date and Approval No: 14/02/2025-315939) for their approval and support in the research process. Their collaboration was essential to the successful completion of this study.

In addition, during the preparation of this manuscript, the authors used an AI-based language tool (ChatGPT by OpenAI) solely to improve the clarity and fluency of the English language. The AI tool was not used for content generation, data analysis, or interpretation. All scientific ideas and findings presented are entirely the work of the authors.

Declaration of Conflicting Interests and Ethics

The authors affirm that there are no known financial, institutional, or personal relationships that could be perceived as a potential conflict of interest with respect to the research, authorship, and publication of this article. The scientific and legal responsibility for manuscripts published in IJATE belongs to the authors. All phases of the research, including design, data collection, analysis, interpretation, and reporting, were conducted independently and without any influence from external parties.

Ethics Committee Number: Ethics Committee of Kırıkkale University (Date and Approval No: 14/02/2025-315939).

Data Availability Statement

All data, measurement instruments, and analysis code used in this research are publicly available through the Open Science Framework (OSF) at <https://osf.io/85q9w/files>

Informed Consent

Written informed consent was obtained from the participants who participated in this study.

Contribution of Authors

M.C.: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Software, Visualization, Writing-original draft, Writing-review & editing. **K.C.:** Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing-original draft, Writing-review & editing. **O.Y.:** Supervision, Methodology, Validation, Writing-review & editing, Project administration.

Orcid

Mehmet Ceylan  <https://orcid.org/0000-0003-0291-4247>

Kevser Ceylan  <https://orcid.org/0009-0004-5674-8189>

Oğuzhan Yoncalık  <https://orcid.org/0000-0003-2289-1257>

REFERENCES































- Bieleke, M., Goetz, T., Yanagida, T., Botes, E., Frenzel, A.C., & Pekrun, R. (2023). Measuring emotions in mathematics: The Achievement Emotions Questionnaire-Mathematics (AEQ-M). *ZDM - Mathematics Education*, 55(2), 269-284. <https://doi.org/10.1007/s11858-022-01425-8>
- Bieleke, M., Gogol, K., Goetz, T., Daniels, L., & Pekrun, R. (2021). The AEQ-S: A short version of the Achievement Emotions Questionnaire. *Contemporary Educational Psychology*, 65, Article 101940. <https://doi.org/10.1016/j.cedpsych.2020.101940>

- Byrne, B.M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (3rd ed.). Routledge.
- Camacho-Morles, J., Slemp, G.R., Pekrun, R., Loderer, K., Hou, H., & Oades, L.G. (2021). Activity achievement emotions and academic performance: A meta-analysis. *Educational Psychology Review*, 33(3), 1051-1095. <https://doi.org/10.1007/s10648-020-09585-3>
- Cañabate, D., Martínez, G., Rodríguez, D., & Colomer, J. (2018). Analysing emotions and social skills in physical education. *Sustainability*, 10(5), Article 1585. <https://doi.org/10.3390/su10051585>
- Ceylan, M., & Karlı, Ü. (2024). Adaptation of the achievement emotions questionnaire for physical education (AEQ- PE) to the Turkish language. *Journal of Sport Sciences Research*, 9(2), 132-148. <https://doi.org/10.25307/jssr.1416025>
- DeVellis, R.F., & Thorpe, C.T. (2021). *Scale development: Theory and applications* (5th ed.). Sage publications.
- Espoz-Lazo, S., Rodríguez Huete, R., Espoz-Lazo, P., Farías-Valenzuela, C., & Valdivia-Moral, P. (2020). Emotional education for the development of primary and secondary school students through physical education: Literature review. *Education Sciences*, 10(8), Article 192. <https://doi.org/10.3390/educsci10080192>
- Fierro-Suero, S., Almagro, B.J., & Sáenz-López, P. (2020). Validation of the achievement emotions questionnaire for physical education (AEQ-PE). *International Journal of Environmental Research and Public Health*, 17(12), 1-12. <https://doi.org/10.3390/ijerph17124560>
- Fornell, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.1177/002224378101800104>
- Fraenkel, J.R., Wallen, N.E., & Hyun, H.H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.
- Garn, A.C., Simonton, K., Dasingert, T., & Simonton, A. (2017). Predicting changes in student engagement in university physical education: Application of control-value theory of achievement emotions. *Psychology of Sport and Exercise*, 29, 93-102. <https://doi.org/10.1016/j.psychsport.2016.12.005>
- Hambleton, R.K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, 1(1), 1-30.
- Hu, L.T., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424. <https://doi.org/10.1037/1082-989X.3.4.424>
- 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), 1-55. <https://doi.org/10.1080/10705519909540118>
- 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(2), 190-201. <https://doi.org/10.1016/j.lindif.2011.04.009>
- Lin, L.C., Huang, P.H., & Weng, L.J. (2017). Selecting path models in SEM: A comparison of model selection criteria. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(6), 855-869. <https://doi.org/10.1080/10705511.2017.1363652>
- Marsh, H.W., Morin, A.J.S., Parker, P.D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, 10, 85-110. <https://doi.org/10.1146/annurev-clinpsy-032813-153700>
- Morin, A.J.S., Arens, A.K., & Marsh, H.W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(1), 116-139. <https://doi.org/10.1080/10705511.2014.961800>
- Morin, A.J.S., Boudrias, J.S., Marsh, H.W., McInerney, D.M., Dagenais-Desmarais, V., Madore, I., & Litalien, D. (2017). Complementary variable- and person-centered approaches to the dimensionality of psychometric constructs: Application to psychological wellbeing at work. *Journal of Business and Psychology*, 32(4), 395-419. <https://doi.org/10.1007/s10869-016-9448-7>

- Muthén, L.K., & Muthén, B.O. (2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.
- Navarro, N., Trigueros, R., González-Bernal, J.J., & Caggiano, V. (2024). A psychometric evaluation of the achievement emotion questionnaire - short version in the context of Italian physical education. *Perceptual and Motor Skills*, 131(5), 1984-2007. <https://doi.org/10.1177/00315125241274820>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315-341. <https://doi.org/10.1007/s10648-006-9029-9>
- Pekrun, R. (2024). Control-value theory: From achievement emotion to a general theory of human emotions. *Educational Psychology Review*, 36(3), 83. <https://doi.org/10.1007/s10648-024-09909-7>
- 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(1), 36-48. <https://doi.org/10.1016/j.cedpsych.2010.10.02>
- Pekrun, R., Marsh, H.W., Elliot, A.J., Stockinger, K., Perry, R.P., Vogl, E., Tilburg, W.A.P. van, Goetz, T., Lüdtke, O., & Vispoel, W.P. (2023). A three-dimensional taxonomy of achievement emotions. *Journal of Personality and Social Psychology: Personality Processes and Individual Differences*, 124(1), 145-178. <https://doi.org/10.1037/pspp0000448.supp>
- Reise, S.P., Bonifay, W.E., & Haviland, M.G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95(2), 129-140. <https://doi.org/10.1080/00223891.2012.725437>
- Pituch, K.A., & Stevens, J.P. (2015). *Applied Multivariate Statistics for the Social Sciences* (3rd ed.). Lawrence Erlbaum. <https://doi.org/10.4324/9781315814919>
- Shao, K., Stockinger, K., Marsh, H.W., & Pekrun, R. (2023). Applying control-value theory for examining multiple emotions in L2 classrooms: Validating the Achievement Emotions Questionnaire – Second Language Learning. *Language Teaching Research*, 94(1), 35-44. <https://doi.org/10.1177/13621688221144497>
- Simonton, K.L., & Garn, A. (2019). Exploring achievement emotions in physical education: The potential for the control-value theory of achievement emotions. *Quest*, 71(4), 434-446. <https://doi.org/10.1080/00336297.2018.1542321>
- Simonton, K.L., & Garn, A.C. (2020). Negative emotions as predictors of behavioral outcomes in middle school physical education. *European Physical Education Review*, 26(4), 764-781. <https://doi.org/10.1177/1356336X19879950>
- Simonton, K.L., Garn, A.C., & Layne, T.E. (2024). An investigation of content-specific unit emotions in secondary physical education. *Education Sciences*, 14(10), Article 1123. <https://doi.org/10.3390/educsci14101123>
- Simonton, K.L., Garn, A.C., & Mercier, K.J. (2023). Expanding the discrete emotions in physical education scale (DEPES): Evaluating emotions with behavior and learning. *Research Quarterly for Exercise and Sport*, 94(1), 35-44. <https://doi.org/10.1080/02701367.2021.1935434>
- Sousa, V.D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, 17(2), 268-274. <https://doi.org/10.1111/j.1365-2753.2010.01434.x>
- Streiner, D.L., Norman, G.R., & Cairney, J. (2024). *Health measurement scales: A practical guide to their development and use* (6th ed.). Oxford University Press. <https://doi.org/10.1093/med/9780192869487.001.0001>
- van Zyl, L.E., & ten Klooster, P.M. (2022). Exploratory Structural Equation Modeling: Practical guidelines and tutorial with a convenient online tool for Mplus. *Frontiers in Psychiatry*, 12, 1-28. <https://doi.org/10.3389/fpsy.2021.795672>
- Wright, P.M., Gordon, B., & Gray, S. (2020). Social and emotional learning in the physical education curriculum. In *Oxford Research Encyclopedia of Education*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.1061>

APPENDIX

Appendix 1. Final form of the AEQ-PE-ES with visual response options.

Enjoyment					
1	I enjoy physical education. <i>Beden eğitimi dersinden zevk alırım.</i>				
2	I look forward to physical education lessons. <i>Beden eğitimi dersini heyecanla beklerim.</i>				
3	Taking part in physical education makes me feel happy. <i>Beden eğitimi dersine katılmak beni mutlu hissettiriyor.</i>				
					
					
	Not at all <i>Hiç</i>	A little <i>Çok az</i>	Somewhat <i>Orta</i>	Quite <i>Oldukça</i>	Very much <i>Çok fazla</i>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anxiety					
4	I feel worried when I think about physical education. <i>Beden eğitimi dersini düşündüğümde endişeleniyorum.</i>				
5	The things I have to do in physical education make me feel worried. <i>Beden eğitimi dersinde yapmam gereken şeyler beni endişelendiriyor.</i>				
6	I am afraid of not being able to do a movement in physical education. <i>Beden eğitimi dersinde bir hareketi yapamamaktan korkarım.</i>				
					
					
	Not at all <i>Hiç</i>	A little <i>Çok az</i>	Somewhat <i>Orta</i>	Quite <i>Oldukça</i>	Very much <i>Çok fazla</i>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Boredom					
7	I find physical education boring. <i>Beden eğitimi dersini sıkıcı bulurum.</i>				
8	Physical education makes me feel overwhelmed. <i>Beden eğitimi dersi beni bunaltıyor.</i>				
9	I feel like time passes slowly in physical education. <i>Beden eğitimi dersinde zamanın yavaş geçtiğini düşünüyorum.</i>				
					
					
	Not at all <i>Hiç</i>	A little <i>Çok az</i>	Somewhat <i>Orta</i>	Quite <i>Oldukça</i>	Very much <i>Çok fazla</i>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>