Finnish students' enjoyment and anxiety levels during fitness testing classes

Finnish students' enjoyment and anxiety levels during fitness testing classes

Mikko Huhtiniemi , Kasper Salin , Jukka Lahti , Arja Sääkslahti , Asko Tolvanen , Anthony Watt & Timo Jaakkola

Published online: 14 Jul 2020.

Submit your article to this journal ⏎

Article views: 528

View related articles ⏎

View Crossmark data ⏎

Routledge
Taylor & Francis Group

Finnish students' enjoyment and anxiety levels during fitness testing classes

Mikko Huhtiniemi [a], Kasper Salin [a], Jukka Lahti[a], Arja Sääkslahti [a], Asko Tolvanen [b], Anthony Watt [c] and Timo Jaakkola [a]

aFaculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland; bFaculty of Education and Psychology, University of Jyväskylä, Jyväskylä, Finland; cCollege of Arts and Education, Victoria University, Melbourne, Australia

ABSTRACT

Background: Fitness testing is a commonly applied learning and teaching practice implemented in both secondary and elementary school physical education (PE). Many teachers believe that by using a variety of different tests, they are able to provide students with feedback regarding their fitness status, and furthermore, increase students' willingness to be physically active later in their lives. However, empirical evidence concerning students' affective responses during fitness testing classes is limited.

Purpose: The primary aim of the study was to investigate whether students' perceptions of enjoyment and anxiety differed between two different types of fitness testing classes and PE in general. In addition, the measurement invariances over time and between Grade 5 (aged 11–12) and Grade 8 (aged 14–15) groups were determined.

Method: A total sample of 645 Finnish Grade 5 (N = 328, 50% boys, mean age = 11.2, SD = 0.36) and Grade 8 students (N = 317, 47% boys, mean age = 14.2, SD = 0.35) participated in the study. Series of multi-group confirmatory factor analyses were conducted to test the level of measurement invariance between general PE and fitness testing classes, and between age groups. Strict factorial invariance was supported for both enjoyment and anxiety scales allowing for latent mean comparisons. Latent mean differences were studied using z-tests.

Results: Grade 5 students perceived significantly lower levels of enjoyment and cognitive processes and a higher level of somatic anxiety in fitness testing classes compared to general PE. Additionally, for Grade 8 students, levels of enjoyment and cognitive processes were significantly lower and somatic anxiety and worry higher in fitness testing classes than in general PE. Furthermore, enjoyment was significantly higher, and cognitive processes, somatic anxiety and worry lower among Grade 5 students compared with Grade 8 students in both contextual PE and during fitness testing class.

Conclusion: Results of this study indicate that students' perceptions of enjoyment were lower in fitness testing classes compared to PE in general. Additionally, students perceived lower levels of cognitive anxiety and higher levels of somatic anxiety in fitness testing classes than in general PE. It is noteworthy that students might not significantly dislike fitness testing per se but instead have significantly more positive affects towards PE in general. Generally, practitioners conducting fitness testing lessons are

---

CONTACT Mikko Huhtiniemi ✉ mikko.huhtiniemi@jyu.fi ▦ Faculty of Sport and Health Sciences, University of Jyväskylä, P.O. Box 35 (L), Jyväskylä 40014, Finland

encouraged to embrace different strategies such as fostering basic psychological needs or promoting mastery climate to facilitate enjoyment and diminish anxiety.

Introduction

Physical education (PE) offers an ideal context for developing students' perceptions and attitudes towards physical fitness, physical performance and physical activity because it reaches the whole age cohort and is implemented by teaching professionals (Sallis et al. 2012). One commonly applied, yet controversial element in school PE is fitness testing (e.g. Simonton, Mercier, and Garn 2019; Jaakkola et al. 2013; Cale and Harris 2009; Silverman, Keating, and Phillips 2008; Rice 2007). According to SHAPE (2017), the fundamental aim of fitness testing is to provide students with necessary knowledge and skills for achieving and maintaining a health-enhancing level of physical activity and fitness. Many teachers consider that by using a variety of different tests they can provide students with feedback regarding their fitness status, and furthermore, increase students' willingness to become or remain physically active (Harris and Cale 2006). Fitness testing has been found to be a positive and enjoyable experience and a useful tool to motivate students towards lifelong physical activity if delivered in an affirming and supportive manner. For example, Jaakkola et al. (2013) found that students had higher perceptions of autonomous motivation in fitness testing sessions than their regular PE classes. Additionally, Goudas, Biddle, and Fox (1994) showed that students with high task orientation and low ego orientation had the highest levels of enjoyment regardless of their results in a 20-m shuttle run test. However, several researchers have reported that fitness testing may lead to negative experiences and cause students to be less interested and involved in PE or general physical activity (e.g. Rice 2007; Naughton, Carlson, and Greene 2006; Corbin 2002). For example, Lodewyk and Muir (2017) demonstrated that Grade 9 girls perceived higher levels of state anxiety and social physique anxiety in fitness testing lessons than in soccer lessons. Furthermore, Hopple and Graham (1995) reported in their qualitative study that Grade 4 and 5 students from the United States had difficulties understanding the purpose of the 1-mile-run test and that they generally had negative perceptions regarding the test. Additionally, in a study by Luke and Sinclair (1991), Canadian Grade 11 boys and girls experienced fitness testing unfavorably and reported it contributing to negative attitudes in PE context. Despite previous research attention, additional investigations on students' affective experiences of fitness testing are needed.

School PE, and fitness testing as part of it, are both examples of contexts where students' positive and negative affective experiences are clearly demonstrated. One example of positive affect is enjoyment which can be characterized as a multidimensional construct closely related to enthusiasm, excitement and perceptions of competence (Hashim, Grove, and Whipp 2008). According to Scanlan and Simons (1992), enjoyment can be verbalized through terms such as 'happiness,' 'liking,' 'pleasure' and 'fun,' and it is therefore seen to represent these more generalized feelings rather than specific emotions such as excitement. Additionally, according to Goetz et al. (2006), enjoyment can be seen as a hierarchically structured concept. This means that one might perceive enjoyment differently in general life than in a specific context such as in PE. Although past studies of enjoyment in fitness testing classes are limited, there is a substantial body of research focused on PE demonstrating consistently high levels of enjoyment among elementary (Carroll and Loumidis 2001; Huhtiniemi et al. 2019) and secondary school students (Soini 2006; Gråsten 2014). Furthermore, enjoyment in PE has been consistently associated with physical activity engagement during school PE (Hashim, Grove, and Whipp 2008; Dishman et al. 2005) and leisure time (Bengoechea et al. 2010; Hashim, Grove, and Whipp 2008; Wallhead and Buckworth 2004).

Although the majority of students find PE enjoyable (Soini 2006), some have reported negative experiences in PE. Barkoukis (2007) proposed that these negative perceptions may rise from

many different factors such as social evaluation, peer comparison or low competence in PE context. The most common negative affect studied in PE is anxiety. More specifically, Barkoukis reported that anxiety in PE classes can be explained through cognitive symptoms (e.g. having negative thoughts), somatic symptoms (e.g. having shortness of breath) and information processing symptoms (e.g. attention disruption). Naturally, many psychosocial factors may trigger responses towards these distinctive manifestations. Factors such as different content areas in PE, class atmosphere, or teachers' interpersonal style may increase or decrease perceptions of anxiety among students. For example, a mastery-oriented motivational climate has been associated with lower anxiety in PE classes (Papaioannou and Kouli 1999; Cecchini et al. 2001). Furthermore, Cox et al. (2011) have demonstrated that social physique anxiety in PE may lead to diminished PE participation and effort. From a broader perspective, there is a plethora of studies focused on test-anxiety in the educational context (Hembree 1988; Zeidner 1998; Von der Embse et al. 2018) demonstrating that test-anxiety associates negatively to a range of behavioral and affective outcomes.

In order to examine how students' affective perceptions vary between fitness testing situations and general PE, we used the concept of different generality levels, previously utilized when studying enjoyment (Goetz et al. 2006) and anxiety (e.g. Zeidner 1998) in educational contexts. One model that has been regularly applied in the context of sport and physical activity is Vallerand's (1997) hierarchical model of intrinsic and extrinsic motivation where it is proposed that motivation and subsequent psychological outcomes such as enjoyment or anxiety occur at three levels, namely global (personality), contextual (life domain) and situational (state). Fitness testing class is an example of a situational level where motivational and affective perceptions arise from the immediate experiences of involvement in a given situation. While the situational level is highly specific, the contextual level represents a more generalized perspective of certain life domains such as education or sport. One example of a contextual level is PE (Jaakkola et al. 2013). The Vallerand's (1997) hierarchical model has been incorporated in multiple studies analyzing students' perceptions in different contexts such as sport (Kowal and Fortier 2000) and PE (Jaakkola et al. 2008). Thus it can be used as a framework to study students' affective perceptions in general PE and in fitness testing situations.

Previous research findings regarding students' affective perceptions in PE fitness testing situations are limited and variable (e.g. Lodewyk and Muir 2017; Jaakkola et al. 2013; Hopple and Graham 1995; Luke and Sinclair 1991). More specifically, a review of the literature reveals that there are no studies examining how students' perceptions of enjoyment and anxiety differ between contextual PE and situational fitness testing class. This knowledge would be useful for PE practitioners and PE teacher educators, and for future intervention development. Therefore, to contribute to the body of knowledge on this topic, the primary aim of this study was to investigate whether students' perceptions of enjoyment and anxiety differed between two different types of fitness testing classes and PE in general. More specifically, the aim was to examine whether students' perceptions of enjoyment, somatic anxiety, cognitive processes and worry differed among PE in general and two fitness testing classes with distinctive content foci (class 1: aerobic endurance, class 2: skills and muscular strength). Furthermore, as the enjoyment and anxiety scales of this study (SCQ-2: Scanlan et al. 2016 and PESAS: Barkoukis et al. 2005) have not been used in fitness testing situations, the additional aim of this study was to investigate the measurement invariance of anxiety and enjoyment scales over time (contextual PE vs. situational fitness testing class 1 and 2) and across groups (Grade 5 and Grade 8 students).

Methods

Participants of the study were 645 Finnish Grade 5 and 8 students recruited from 36 classes and 12 schools in the Southern, Western, and Central regions of Finland. Invitations to participate were sent to schools in different regions and those willing to participate were recruited for the study. Schools represented both urban and rural areas, and followed the national core curriculum with no optional study lines (e.g. sport or math emphasis). Teachers in Grade 8 were all specialized in PE whereas in Grade 5 teachers were generalist class teachers with basic training in PE. The sample of Grade 5 students included

164 boys and 164 girls with an average age of 11.2 years (SD = .36) and the sample of Grade 8 students included 150 boys and 167 girls with an average age of 14.2 years (SD = .35). Students with disabilities or special education needs did not participate in the study. Therefore, the sample comprised students who were engaged in the regular school program and following the national curriculum.

Procedure

Before commencing the data collection, the ethics committee of the local university approved the study protocol. Informed consent forms from both students and their guardians were obtained prior to the study. Participation was voluntary and students had the opportunity to withdraw from the study at any time. The first questionnaire, assessing contextual perceptions of PE, was administered in September before a regular PE class (T0). Students were asked to think about their general experiences of PE. The second questionnaire was completed two weeks later immediately after conclusion of the first fitness testing lesson (T1). This time, students were specifically asked to reflect upon their perceptions of the fitness testing lesson. Finally, the third questionnaire was completed one week later, immediately after the second fitness testing lesson (T2). Again, students were asked to consider their perceptions of the lesson they just concluded. During the questionnaires, students were allowed to ask for guidance if they did not comprehend some of the questions. Questionnaires were administered by trained PE teachers briefed regarding the research who followed written step-by-step instructions during the procedure. To further corroborate the reliability, a pilot of the procedure was conducted prior to the commencement of the main study. In the pilot, no problems were encountered with the protocol.

The first fitness testing class (90 min) comprised of 20 meters shuttle run test (20mSRT; Léger et al. 1988) (aerobic endurance and movement skills) and flexibility tests including: squat (flexibility of the pelvis and lower limbs), lower back extension (range of motion of the lower back and hip area joints) and flexibility of the right and left shoulders (flexibility of upper limbs and shoulder area). The second fitness testing class (90 min) included four tests: curl-ups (abdominal strength and endurance), push-ups (upper body strength), 5-leaps (lower limb strength, speed, dynamic balance skills and movement skills) and throwing-catching combination (object control skills, perceptual motor skills and upper limb strength) (Jaakkola et al. 2012). The fitness tests and accompanying protocol were obtained from the Finnish Move!® system for monitoring physical functioning capacity. (www.edu.fi/move). The intended learning outcomes for the fitness testing lessons were also obtained from the national guidelines of the Move! system. According to those guidelines, the overall goal of the Move! fitness testing is to provide students with information concerning their physical fitness, and to encourage them to independently take care of their physical functioning capacity. It is noteworthy, that the Finnish PE curriculum forbids using the fitness testing results as a basis for grading.

Contextual measures

Enjoyment in physical education. The Finnish version of the Enjoyment subscale from the Sport Commitment Questionnaire -2 (SCQ-2; Scanlan et al. 2016) was used to analyze enjoyment in PE. The Scale comprises five items (e.g. 'Physical education is fun') which are rated on a 5-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. Respondents were asked to think about their overall PE experiences. The validity and reliability of the enjoyment scale has been previously reported when used with Finnish Grade 5 (aged 11–12) (CFI = 1.00, RMSEA = 0.10, Cronbach's alpha = 0.91) and Grade 8 (aged 14–15) (CFI = 1.00, RMSEA = 0.07, Cronbach's alpha = 0.95) students during PE classes (Huhtiniemi et al. 2019).

Anxiety in physical education. The Finnish version of the Physical Education State Anxiety Scale (PESAS; Barkoukis et al. 2005) was used to measure anxiety in PE. The scale assesses three dimensions of anxiety, namely somatic anxiety, cognitive processes and worry. Each dimension is formed from six items rated on a 5-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. Somatic anxiety refers to perceptions of physical symptoms (e.g. 'I have a sense of pressure

on my chest'), cognitive processes refer to symptoms related to information processing, such as memory and attention, during the activity (e.g. 'I find it difficult to focus on the task presented'), and worry refers to negative expectations from involving in the activity and the consequences of possible failures (e.g. 'I'm afraid of making mistakes while performing the exercises'). PESAS has been used in the past with Finnish students and demonstrated adequate psychometric properties (CFI = 0.94; RMSEA = 0.05; Cronbach alphas between 0.73 and 0.88) (Liukkonen et al. 2010).

Situational measures during fitness testing classes

Enjoyment during fitness testing classes was measured using the same SCQ-2 enjoyment subscale as described earlier with the exception of different situational specific wording on items (e.g. 'This fitness testing class was fun' versus 'physical education is fun'). These revisions were made to emphasize the change from contextual PE to fitness testing situation. Similarly, anxiety during fitness testing class was measured using the PESAS scale with the exception of using a different stem ('During this fitness testing class … ') emphasizing the change from contextual PE to fitness testing situation.

Statistical analyses

Before proceeding to the main analyses, the data were screened for outliers (values below or above the possible range of 1–5) and missing data patterns. All values were inside the range (1–5). Statistical analyses were performed, and missing data handled, using Mplus version 8.2 (Muthén and Muthén 2017) and implementing the robust full-information maximum-likelihood (MLR) estimation method. A series of multi-group confirmatory factor analyses (CFA) were conducted to test the measurement invariance across groups (Grade 5 and Grade 8) and over time (T0–T2). Measurement invariance (i.e. does the scale function in a similar way over time and across groups?) is a precondition for investigating latent mean differences among the study variables and it involves incorporating increasingly stringent steps of constraining different model parameters (Kline 2015). The overall model fit was evaluated using multiple indicators, as suggested by Ntoumanis and Myers (2016). More specifically, the chi-square goodness-of-fit statistics (χ^2), the comparative fit index (CFI), the Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) were used. To interpret these indices, we followed previously recommended guidelines. For CFI and TLI, cut-off values close to .95, for RMSEA values lower than .06, and for SRMR values lower than .08, were considered good (Hu and Bentler 1999).

Invariance of factor loadings, item intercepts and residuals over time and across grade-level groups were tested using the following hierarchically constructed procedure: Model M1, all parameters freely estimated over measurement time and across Grade level groups; Model M2, factor loadings set equal over time separately in both Grade level groups; Model M3, factor loadings and intercepts set equal over time separately in both Grade level groups; Model M4, factor loadings, intercepts and residual variances of observed variables set equal over time separately in both Grade level groups; Model M5, factor loadings, intercepts, and residual variances of observed variables set equal over time and factor loadings set equal between Grade level groups; Model M6, factor loadings, intercepts, and residual variances of observed variables set equal over time and factor loadings and intercepts set equal between Grade level groups; and Model M7, factor loadings, intercepts and residual variances of observed variables set equal over time and between Grade level groups. In all models, indicator-specific effects over time were accounted for by allowing autocorrelation (Little 2013).

When evaluating the level of measurement invariance between nested models, we used the Satorra–Bentler-corrected chi-square difference test (S–Bχ^2) along with the RMSEA (Satorra and Bentler 2001; Hu and Bentler 1999). A non-significant change in the S–Bχ^2 indicates that the invariance holds when comparing the more constrained model to the less constrained model. However, the χ^2 value has been recognized as overly sensitive to large sample sizes (Cheung and Rensvold

2009). Therefore, in the case of a significant χ^2 difference test, we evaluated the amount of difference between the nested models by comparing RMSEA values. When comparing the models, a criterion of ΔRMSEA $\leq$ .015 was seen acceptable (Chen 2007).

After evaluating the measurement invariance, we proceeded by investigating the mean differences in the latent constructs. With repeated measures, the mean of a latent variable is usually set to 0 in one group or time point (as a reference) and freely estimated in other groups or time points (Muthén and Muthén 2017). However, in this study we incorporated a model constraint setting the sum of item intercepts over time and across groups to 0 in order to estimate the latent means on the same scale as the original items (1–5). Therefore, the interpretation of the mean levels, as well as the mean differences were more convenient. Statistical significance of mean differences was determined using z-tests. Finally, 95% confidence interval levels and effect sizes using Cohen's d were calculated.

Results

Preliminary analyses

The analysis process commenced by establishing configural baseline models (model M1; see Table 1) for the three subscales of anxiety and for the enjoyment. Based on the model fit criteria by Hu and Bentler (1999), all models demonstrated a good fit. The sequential models and different levels of invariance were tested in two waves. We first investigated the invariance of factor loadings (metric or weak factorial invariance), then the equality of item intercepts (scalar or strong factorial invariance) and finally equalities of item residuals (error or strict invariance) over time separately between the two groups (models M2–M4). We continued by applying the same parameter constraints over time and also across both groups (models M5–M7). The fit indices for all nested models of enjoyment and anxiety subscales are presented in Table 1 and standardized item loadings obtained from model M7 are presented in Table 2.

For enjoyment, results indicated that all levels of measurement invariance (i.e. weak, strong, and strict factorial invariance) held over time and across groups. This is shown by the small changes in model fit (ΔRMSEA $\leq$.004) when comparing more constrained models to the more freely estimated models. Similarly, for cognitive processes, somatic anxiety, and worry, all levels of measurement invariances were supported based on the small changes in model fit. The values for ΔRMSEA were $\leq$.014, $\leq$.006 and $\leq$.005, respectively.

Main analyses

After establishing the measurement invariance, we proceeded by investigating the mean differences of latent constructs. As presented in Table 3, three pairwise comparisons were made over time for both Grade 5 and Grade 8 groups. For Grade 5 students, results showed that enjoyment was significantly lower in fitness testing classes (T1 and T2) compared to PE in general (T0). There was no difference between the two fitness testing classes (T1 vs. T2) in enjoyment. Level of cognitive processes was significantly lower and level of somatic anxiety higher in fitness testing classes than in PE in general. When comparing the two fitness testing classes, there were no statistically significant differences between the groups in levels of cognitive processes. Also, results showed that somatic anxiety was higher on the first fitness testing class (T1) compared to the second fitness testing class (T2). Finally, there were no statistically significant differences in worry among the three time points.

For Grade 8 students, results indicated that levels of enjoyment were significantly lower in fitness testing classes (T1 and T2) compared to PE in general (T0). There was no difference in enjoyment between the two fitness testing classes (T1 vs. T2). According to the results for cognitive processes, there was a statistically significant difference between PE in general (T0) and the first fitness testing

Table 1. Measurement invariance of cognitive processes, somatic anxiety, worry, and enjoyment over time and across grade level.

Scale	χ^2	df	p-Value	RMSEA	CFI	TLI	SRMR	$\Delta\chi^2$	Δdf	p-Value
Cognitive processes M1	265.75	222	0.024	0.025	0.984	0.978	0.045	–	–	–
Cognitive processes M2	287.04	242	0.025	0.024	0.984	0.979	0.052	20.81	20	0.408
Cognitive processes M3	382.66	262	0.000	0.038	0.956	0.949	0.057	117.63	20	0.000
Cognitive processes M4	404.51	286	0.000	0.036	0.957	0.954	0.063	27.95	24	0.262
Cognitive processes M5	407.77	291	0.000	0.035	0.957	0.955	0.065	2.75	5	0.738
Cognitive processes M6	424.30	296	0.000	0.037	0.953	0.952	0.065	20.62	5	0.001
Cognitive processes M7	441.72	302	0.000	0.038	0.949	0.948	0.068	13.00	6	0.043
Somatic anxiety M1	323.16	228	0.000	0.036	0.973	0.964	0.050	–	–	–
Somatic anxiety M2	357.45	248	0.000	0.037	0.970	0.962	0.057	34.68	20	0.022
Somatic anxiety M3	459.05	268	0.000	0.047	0.947	0.939	0.063	111.13	20	0.000
Somatic anxiety M4	555.89	292	0.000	0.053	0.926	0.923	0.072	81.54	24	0.000
Somatic anxiety M5	564.21	297	0.000	0.053	0.926	0.923	0.072	8.25	5	0.143
Somatic anxiety M6	583.99	302	0.000	0.054	0.921	0.920	0.073	21.28	5	0.001
Somatic anxiety M7	628.10	308	0.000	0.057	0.911	0.911	0.080	32.94	6	0.000
Worry M1	282.64	228	0.008	0.027	0.989	0.985	0.036	–	-	–
Worry M2	297.65	248	0.017	0.025	0.990	0.987	0.039	11.36	20	0.936
Worry M3	347.02	268	0.001	0.030	0.983	0.981	0.041	58.51	20	0.000
Worry M4	374.63	292	0.001	0.030	0.983	0.982	0.044	28.59	24	0.236
Worry M5	381.12	297	0.001	0.030	0.982	0.982	0.046	6.51	5	0.260
Worry M6	395.22	302	0.000	0.031	0.980	0.980	0.046	17.02	5	0.004
Worry M7	396.53	308	0.001	0.030	0.981	0.981	0.047	4.18	6	0.652
Enjoyment M1	184.19	142	0.010	0.031	0.992	0.987	0.036			
Enjoyment M2	219.99	158	0.001	0.035	0.988	0.983	0.050	44.87	16	0.000
Enjoyment M3	256.16	174	0.000	0.038	0.983	0.980	0.049	39.34	16	0.001
Enjoyment M4	302.61	194	0.000	0.042	0.978	0.976	0.058	38.56	20	0.008
Enjoyment M5	312.93	198	0.000	0.043	0.977	0.975	0.060	13.12	4	0.011
Enjoyment M6	321.76	202	0.000	0.043	0.976	0.975	0.061	9.89	4	0.042
Enjoyment M7	319.28	207	0.000	0.041	0.977	0.977	0.061	2.53	5	0.772

M1: Configural – all parameters are freely estimated across measurement time and across school classes, M2: Factor loadings are set equal across measurement time separately in both school classes, M3: Factor loadings and intercepts are set equal across measurement time separately in both school classes, M4: Factor loadings, intercepts and residual variances of observed variables are set equal across measurement time separately in both school classes, M5: Factor loadings, intercepts and residual variances of observed variables are set equal across measurement time and factor loadings are set equal between school classes, M6: Factor loadings, intercepts and residual variances of observed variables are set equal across measurement time and factor loadings and intercepts are set equal between school classes, M7: Factor loadings, intercepts and residual variances of observed variables are set equal across measurement time and school classes. The final three columns describe χ^2 differences between nested models.

class (T1: aerobic endurance) but not between PE in general (T0) and the second fitness testing class (T2: skill and strength), or between the two fitness testing classes (T1 vs. T2). The level of somatic anxiety was significantly higher in first fitness testing class (T1), and also in second fitness testing class (T2) compared to PE in general (T0). In addition, somatic anxiety was significantly higher in T1 compared to T2. Finally, there was a statistically significant, but weak, difference between the level of worry on T0 and T1. Levels of worry did not differ between T0 and T2 or between T1 and T2.

Differences between Grade 5 and Grade 8 students in study variables were also analyzed at each measurement point. As can be seen from Table 4, results indicated that levels of enjoyment were significantly higher among Grade 5 students at all three time points. Furthermore, the mean levels of cognitive processes and worry were lower among Grade 5 students than Grade 8 students at all different time points. For somatic anxiety, there was a statistically significant difference at T0 and T1 showing lower mean scores for Grade 5 students, but no significant difference at T2 was found.

Discussion

This study aimed to investigate whether students' perceptions of enjoyment and anxiety differed between two different types of fitness testing classes and PE in general. This was the first study to

Table 2. Standardized item loadings of the multi-group confirmatory factor analyses.

	Grade 5			Grade 8		
	PE in general (T0) Estimate	Fitness testing class 1 (T1)	Fitness testing class 2 (T2)	PE in general (T0) Estimate	Fitness testing class 1 (T1)	Fitness testing class 2 (T2)
Item	(S.E.)	Estimate (S.E.)	Estimate (S.E.)	(S.E.)	Estimate (S.E.)	Estimate (S.E.)
Enjoyment						
e1	0.82 (.020)	0.89 (.011)	0.90 (.011)	0.86 (.014)	0.86 (.014)	0.85 (.016)
e2	0.85 (.018)	0.91 (.010)	0.92 (.010)	0.89 (.013)	0.89 (.012)	0.88 (.015)
e3	0.81 (.020)	0.88 (.010)	0.89 (.010)	0.85 (.013)	0.85 (.012)	0.84 (.014)
e4	0.87 (.015)	0.92 (.008)	0.93 (.008)	0.90 (.011)	0.91 (.010)	0.90 (.012)
e5	0.85 (.017)	0.91 (.009)	0.92 (.008)	0.89 (.012)	0.89 (.011)	0.88 (.012)
Cognitive processes						
a1	0.60 (.030)	0.61 (.035)	0.61 (.033)	0.67 (.031)	0.70 (.035)	0.70 (.032)
a2	0.72 (.028)	0.72 (.031)	0.73 (.030)	0.78 (.022)	0.80 (.024)	0.80 (.022)
a3	0.68 (.028)	0.69 (.033)	0.70 (.031)	0.75 (.026)	0.77 (.029)	0.77 (.026)
a4	0.65 (.032)	0.66 (.035)	0.66 (.033)	0.72 (.028)	0.74 (.032)	0.74 (.030)
a5	0.69 (.028)	0.70 (.031)	0.70 (.029)	0.76 (.025)	0.78 (.027)	0.78 (.025)
a6	0.67 (.032)	0.68 (.034)	0.68 (.032)	0.74 (.025)	0.76 (.026)	0.76 (.025)
Somatic anxiety						
a7	0.69 (.030)	0.80 (.020)	0.79 (.020)	0.75 (.025)	0.81 (.018)	0.78 (.023)
a8	0.69 (.031)	0.80 (.020)	0.79 (.022)	0.75 (.024)	0.81 (.018)	0.78 (.023)
a9	0.59 (.037)	0.72 (.028)	0.71 (.032)	0.66 (.034)	0.73 (.027)	0.70 (.034)
a10	0.68 (.029)	0.79 (.021)	0.78 (.023)	0.74 (.028)	0.80 (.020)	0.77 (.025)
a11	0.67 (.031)	0.78 (.023)	0.77 (.026)	0.73 (.031)	0.79 (.022)	0.76 (.027)
a12	0.36 (.030)	0.48 (.030)	0.47 (.031)	0.42 (.029)	0.49 (.027)	0.46 (.029)
Worry						
a13	0.79 (.020)	0.82 (.017)	0.83 (.016)	0.82 (.016)	0.83 (.015)	0.83 (.016)
a14	0.79 (.019)	0.82 (.017)	0.84 (.015)	0.83 (.018)	0.84 (.017)	0.83 (.019)
a15	0.65 (.027)	0.69 (.027)	0.71 (.024)	0.70 (.028)	0.71 (.027)	0.70 (.028)
a16	0.85 (.016)	0.88 (.013)	0.89 (.012)	0.88 (.012)	0.89 (.011)	0.88 (.012)
a17	0.79 (.021)	0.81 (.018)	0.83 (.016)	0.82 (.017)	0.83 (.015)	0.83 (.017)
a18	0.80 (.017)	0.83 (.016)	0.85 (.015)	0.84 (.015)	0.85 (.015)	0.84 (.017)

investigate the measurement invariance of SCQ-2 Enjoyment subscale and PESAS over time and between Grade 5 and Grade 8 students. Additionally, this was the first study to investigate whether students' perceptions of enjoyment and anxiety differ between PE in general and fitness testing classes.

Enjoyment in fitness testing and generally in PE

Results showed that both Grade 5 and Grade 8 students perceived lower levels of enjoyment in fitness testing classes than in general PE. As such, it indicates that fitness testing as a content area in PE might be generating less feelings of pleasure, fun and liking (Scanlan et al. 2016) among elementary and secondary school students than PE in general. However, it is noteworthy that on average students still perceived moderate levels of enjoyment in fitness testing classes. In line with previous studies (Carroll and Loumidis 2001; Soini 2006; Gråsten 2014; Huhtiniemi et al. 2019), levels of enjoyment towards PE in general were relatively high (see Table 3), indicating that students may not specifically dislike fitness testing but have significantly more positive feelings towards PE in general. Nonetheless, deflated levels of enjoyment in fitness testing situations might cause students to be less engaged towards fitness testing or fitness development in PE. Previous intervention studies in PE have shown that students' enjoyment can be positively influenced by emphasizing effort, learning, co-operation and personal development (Barkoukis, Tsorbatzoudis, and Grouios 2008). Additionally, studies have shown that PE enjoyment is positively associated with higher perceptions of mastery climate and basic psychological needs of autonomy, competence, and relatedness (Cox, Smith,

Table 3. Mean differences of study variables among PE in general and fitness testing classes.

Grade 5	T0* M (SD)	T1** M (SD)	T2*** M (SD)	T1 vs. T0		T2 vs. T0		T1 vs. T2	
				p (z-test)	Cohen d (95% CI)	*p* (z-test)	Cohen d (95% CI)	*p* (z-test)	Cohen d (95% CI)
Enjoyment	4.136 (0.897)	3.369 (1.205)	3.382 (1.259)	.000	−0.856 (−1.008, −0.705)	.000	−0.841 (−0.997, −0.685)	.749	−0.015 (−0.094, 0.063)
Cognitive processes	1.591 (0.570)	1.444 (0.583)	1.448 (0.588)	.000	−0.258 (−0.360, −0.156	.000	−0.250 (−0.357, −0.143)	.890	−0.008 (−0.099, 0.084)
Somatic anxiety	2.224 (0.780)	2.743 (1.084)	2.547 (1.057)	.000	0.666 (0.512, 0.820)	.000	0.414 (0.277, 0.552)	.000	0.252 (0.136, 0.368)
Worry	2.117 (0.922)	2.139 (1.021)	2.146 (1.097)	.621	0.024 (−0.056, 0.104)	.538	0.032 (−0.054, 0.118)	.857	−0.008 (−0.083, 0.067)
Grade 8									
Enjoyment	3.519 (1.052)	2.485 (1.062)	2.574 (1.000)	.000	−0.983 (−1.124, −0.841)	.000	−0.898 (−1.035, −0.761)	.105	0.085 (−0.001, 0.170)
Cognitive processes	1.808 (0.686)	1.681 (0.731)	1.766 (0.744)	.012	−0.184 (−0.305, −0.064)	.393	−0.060 (−0.176, 0.055)	.054	−0.124 (−0.231, −0.017)
Somatic anxiety	2.430 (0.932)	2.957 (1.118)	2.618 (1.024)	.000	0.565 (0.420, 0.710)	.004	0.201 (0.080, 0.322)	.000	0.364 (0.236, 0.492)
Worry	2.344 (1.051)	2.466 (1.097)	2.432 (1.062)	.045	0.116 (0.020, 0.213)	.146	0.084 (−0.012, 0.180)	.538	0.033 (−0.054, 0.119)

*T0 = contextual PE.
**T1 = fitness testing class 1 (20 meters shuttle run test, mobility).
***T2 = fitness testing class 2 (curl-ups, push-ups, catching-throwing combination, 5-leaps).

Table 4. Mean differences of study variables between Grade 5 and Grade 8 students.

	T0* (Grade 5 vs. Grade 8)		T1** (Grade 5 vs. Grade 8)		T2*** (Grade 5 vs. Grade 8)	
	p (z-test)	Cohen d (95% CI)	*p* (z-test)	Cohen d (95% CI)	*p* (z-test)	Cohen d (95% CI)
Enjoyment	0.000	0.632 (0.486, 0.777)	0.000	0.904 (0.730, 1.077)	0.000	0.827 (0.651, 1.003)
Cognitive processes	0.000	−0.344 (−0.491, −0.198)	0.000	−0.377 (−0.536, −0.218)	0.000	−0.505 (−0.673, −0.336)
Somatic anxiety	0.007	−0.240 (−0.386, −0.095)	0.030	−0.249 (−0.436, −0.061)	0.466	−0.082 (−0.268, 0.103)
Worry	0.006	−0.230 (−0.369, −0.091)	0.000	−0.332 (−0.481, −0.182)	0.003	−0.289 (−0.448, −0.131)

*T0 = contextual PE.
**T1 = fitness testing class 1 (20 meters shuttle run test, flexibility).
***T2 = fitness testing class 2 (curl-ups, push-ups, catching-throwing combination, 5-leaps).

and Williams 2008; Ommundsen and Kvalø 2007). Although there is a lack of intervention studies investigating specifically fitness testing situations it is likely that enjoyment during testing classes can be promoted through implementing similar strategies to those used in general PE. It should also be noted that fitness testing classes might inherently contain undesirable social or behavioral factors such as peer comparison, norm-referencing or diminished opportunities for autonomous behavior that undermine feelings of enjoyment.

Interestingly, there were no differences in enjoyment levels between the two fitness testing classes although the classes consisted of different types of tests. The first fitness testing class included a 20 meters shuttle run test which can be perceived as strenuous and unpleasant because it requires working near maximal aerobic capacity (Silverman, Keating, and Phillips 2008). In contrast, the second fitness testing class included more skill-related measures and muscular strength measures. For example, the throwing-catching combination (Jaakkola et al. 2012) where one repeatedly throws a tennis ball to the wall and catches it after a bounce could easily be perceived as a fun activity that students might want to do during recess or free time. Yet, despite the content of the testing classes clearly differing, perceived enjoyment remained relatively stable between different test situations for both Grade 5 and Grade 8 students. This pattern might exist because the two fitness testing lessons share similar elements related to pedagogical aspects such as teachers' teaching style and chosen didactic approach. For example, it is likely that teachers used same kind of teaching style techniques while giving instructions or feedback during both fitness testing classes which is indicative of the style of teaching rather than the content of the fitness testing as being more influential in mediating the students' experiences of enjoyment.

Results indicated that Grade 5 students had higher enjoyment ratings than Grade 8 students at all three time points. This reflects previous studies that have revealed declining trends on individual level in students' general PE enjoyment (Yli-Piipari et al. 2012; Barkoukis, Ntoumanis, and Thøgersen-Ntoumani 2010). For example, Barkoukis, Ntoumanis, and Thøgersen-Ntoumani (2010) reported the enjoyment levels of 12-year-old Greek students declined across a 3-year period. During this age period, students go through several biological, social and psychological changes that might negatively affect their self-confidence, and therefore, reduce their feelings of enjoyment. In addition, it might be that the associated learning goals of fitness testing are not clear or adequately suited for the needs of older students which might cause deflated feelings of enjoyment.

Anxiety in fitness testing situations and generally in PE

For both Grade 5 and 8 students, levels of cognitive processes were lower in fitness testing classes than in general PE. In other words, cognitive processes were not stimulated to the same extent in fitness testing situations as in general PE classes. As cognitive processes dimension refers to symptoms related to information processing and cognitive reactions during the activity (Barkoukis, Tsorbatzoudis, and Grouios 2008; Barkoukis et al. 2005; Schwarzer 1986), it might be that these reactions

are typically present during general PE classes rather than in fitness testing classes where the activities performed are very precise. Additionally, lower levels of the cognitive processes dimension might be an outcome of the more structured and teacher-oriented approach during fitness testing lessons where students have less opportunities to make choices or to be creative. Furthermore, the lack of difference in cognitive processes between the two fitness testing classes may also be due to the highly structured and teacher-oriented lessons. Analysis of the age group differences revealed that Grade 8 students experienced higher levels in the cognitive processes dimension both in general PE and in fitness testing classes than Grade 5 students. This may be due to Grade 8 students' maturation level being mediated by their biological, social and cognitive development during the early adolescence years. At this phase of puberty, adolescents tend to have stronger feelings related to self-conscious emotions such as anxiety (Wigfield, Lutz, and Wagner 2005; Eccles and Roeser 2011).

Somatic anxiety, in contrast to cognitive processes, was higher in both fitness testing classes than in general PE for Grade 5 and 8 students. This was not entirely surprising, as the fitness testing situations and physical tests encourage students to perform near their maximal physical capacity. Also, as somatic anxiety has been shown to have a curvilinear relationship with performance (i.e. moderate levels lead to optimum performance) (Craft et al. 2003), it is logical that elevated levels of somatic anxiety occurred in the performance-related fitness testing situation rather than in general PE. Moreover, as somatic anxiety captures phenomena like 'shortness of breath', 'discomfort while breathing', 'feeling dizzy' and 'feeling as if something is choking one' (Barkoukis et al. 2005), it is reasonable to see higher somatic anxiety in the first fitness testing class which included 20mSRT than in the second testing class which included skill and strength related tests. While interpreting these results, it should be noted that somatic anxiety symptoms are linked to body's normal reactions to physical exertion.

In general PE, Grade 5 students' somatic anxiety levels were lower when compared to Grade 8 students. This may be due in part to the elementary school PE curriculum, which places no emphasis on the physical intensity levels of the students and is generally more play-oriented than secondary school PE curriculum (Finnish National Board of Education 2014). Conversely, physical demands in Grade 8 PE curriculum are typically higher which could lead to more pronounced somatic anxiety levels compared to Grade 5 students. Another reason for lower somatic anxiety in Grade 5 could be that in Finland, Grades 1–6 are predominantly taught by generalist classroom teachers with only basic qualifications in PE whereas Grades 7–9 are taught by specialist PE teachers. Therefore, specific teaching qualifications in PE might lead to more active and demanding lessons (Telford et al. 2013), and as a consequence, promote an increase in the occurrence of somatic symptoms such as feeling dizziness or sense of pressure in the chest. As previously mentioned, somatic anxiety symptoms are related to body's normal reactions to physical exertion. Therefore, elevated levels of somatic anxiety are not necessarily negative if they occur during physically demanding activity. Also, other aspects of anxiety should be simultaneously considered to more comprehensively understand one's experiences.

Levels of worry among Grade 5 and Grade 8 students were approximately the same in general PE and during fitness classes, except for the slightly higher value for Grade 8 students in the first testing class. Overall, and contrary to previous findings (e.g. Hopple and Graham 1995; Luke and Sinclair 1991), the pattern of worry level indicates that students do not necessarily have inflated negative expectations of the fitness testing activities or that their fear of getting low results or performing poorly (Barkoukis 2007) is not increased. Earlier studies have shown that worry is a negative predictor of performance in physical activity contexts, such as sport and school PE (Barkoukis et al. 2005; Woodman and Hardy 2003). In considering this pattern, it could be seen as a positive outcome that fitness testing provokes no more worry than general PE. Additionally, levels of worry did not differ between the two fitness testing classes in either of the grade-level groups indicating that different test batteries did not seem to provoke increased achievement pressures.

Grade 5 students perceived lower levels of worry both in general PE and in fitness testing situations compared to Grade 8 students. Again, this might be an age-related issue as Grade 8 students are at, or are getting towards, puberty which could amplify their anxiety levels during physical

activity, especially when undertaken with peers. Another explanation for Grade 8 students' higher level of worry could stem from the PE assessment or numeric grading which is usually introduced to Finnish students from Grade 7 onwards (Finnish National Board of Education 2014). However, it should be noted that the current PE curriculum in Finland – that includes the new national fitness monitoring system – clearly forbids using fitness test results as a basis for grading in PE (Salin and Huhtiniemi 2018).

Measurement invariance

Results indicated that enjoyment and all three anxiety subscales possessed full factorial invariance between different settings and across age groups. It should be acknowledged that full measurement invariance is rarely achieved in most empirical studies (Van De Schoot et al. 2015). This demonstrates that Grade 5 and Grade 8 students perceived the enjoyment and anxiety scales in a similar way when answering questions concerning PE in general (contextual) and answering questions after fitness testing classes (situational). In general, these findings indicate that comparisons between the study variables in different settings and age groups are plausible. However, caution is still warranted especially from a cross-cultural perspective as, for example, instrument adaptation and translation process or social desirability can cause variation in how respondents perceive the questions (Davidov et al. 2014).

Study limitations and future research

The current study has several limitations that need to be considered while interpreting the results. First, the study sample was not randomly selected which limits the representativeness of the results. Second, as students were asked to complete the three questionnaires in a relatively short timeframe, a fatigue or question-order effect (Pustejovsky and Spillane 2009) could influence their responses. Yet, the invariance analysis showed that students perceived the scales in a similar way at each time point. Third, although the study protocol was carefully presented to the teachers, there is no way of knowing how well they followed the written instructions as there were no observations or recordings of the classes available. By adding an objective measure, such as video or voice recording to assess teachers' actions during the classes would therefore increase the reliability of the study. However, adding an external person or device to the situation might significantly affect students' behavior, physical performance and cognitive perceptions.

In the future, it would be interesting to see intervention studies aiming to change students' affective experiences during fitness testing classes, and to test whether changes at the situational level have an effect at the contextual level, or vice versa. These bottom-up or top-down effects have been previously studied in sport contexts targeting motivational constructs (e.g. Kowal and Fortier 2000) but not in the physical education fitness testing context. It would also be interesting to study whether different subgroups of students (e.g. based on gender, ethnicity or special needs) perceive enjoyment and anxiety differently during fitness testing lessons. Additionally, as previous research has shown that PE may affect PA participation during leisure time (Hagger et al. 2003), it would be valuable to investigate how enjoyment and anxiety in school fitness testing classes impact students' willingness for fitness development and physical activity outside school.

Conclusions and practical implications

The findings of this study indicate that students' perceptions of enjoyment and anxiety towards fitness testing classes differ from their perceptions towards PE in general. More specifically, Grade 5 and Grade 8 students' perceptions of enjoyment were lower in fitness testing classes compared to PE in general. Additionally, students perceived lower levels of cognitive anxiety and higher levels of somatic anxiety in fitness testing classes than in general PE. Levels of worry among Grade 5 and

Grade 8 students were approximately at the same level in general PE and during fitness testing classes, except for the slightly higher value for Grade 8 students in the first testing class. Practitioners can make use of the results while planning and conducting fitness testing sessions for different aged children. Although the reasons underpinning affective experiences during fitness testing were not investigated in this study, previous intervention studies have shown that PE enjoyment is positively associated with higher perceptions of mastery climate and basic psychological needs of autonomy, competence and relatedness (Cox, Smith, and Williams 2008; Ommundsen and Kvalø 2007). Previous research has also shown that mastery climate is associated with lower levels of anxiety in PE (Papaioannou and Kouli 1999; Cecchini et al. 2001) and that state anxiety is negatively linked to enjoyment in the PE context (Yli-Piipari et al. 2009). Therefore, adopting strategies that promote mastery climate and need fulfillment are recommended to increase enjoyment and reduce anxiety in PE and also in fitness testing classes.

Disclosure statement

Funding

ORCID

Mikko Huhtiniemi http://orcid.org/0000-0002-1643-2671
Kasper Salin http://orcid.org/0000-0001-6937-4535
Arja Sääkslahti http://orcid.org/0000-0003-4354-0990
Asko Tolvanen http://orcid.org/0000-0001-6430-8897
Anthony Watt http://orcid.org/0000-0002-1084-750X
Timo Jaakkola http://orcid.org/0000-0003-1227-2460

References

Barkoukis, V. 2007. "Experience of State Anxiety in Physical Education." In *Psychology for Physical Educators: Student in Focus*, edited by J. Liukkonen, Y. Auweele, D. Alfermann, B. Vereijken, and Y. Theodorakis, 57–72. Champaign, IL: Human Kinetics.

Barkoukis, V., N. Ntoumanis, and C. Thøgersen-Ntoumani. 2010. "Developmental Changes in Achievement Motivation and Affect in Physical Education: Growth Trajectories and Demographic Differences." *Psychology of Sport and Exercise* 11 (2): 83–90.

Barkoukis, V., H. Tsorbatzoudis, and G. Grouios. 2008. "Manipulation of Motivational Climate in Physical Education: Effects of a 7-Month Intervention." *European Physical Education Review* 14: 373–394.

Barkoukis, V., H. Tsorbatzoudis, G. Grouios, and A. Rodafinos. 2005. "The Development of a Physical Education State Anxiety Scale: a Preliminary Study." *Perceptual and Motor Skills* 100 (1): 118–128.

Bengoechea, E. G., C. M. Sabiston, R. Ahmed, and M. Farnoush. 2010. "Exploring Links to Unorganized and Organized Physical Activity During Adolescence: The Role of Gender, Socioeconomic Status, Weight Status, and Enjoyment of Physical Education." *Research Quarterly for Exercise and Sport* 81: 7–16.

Cale, L., and J. Harris. 2009. "Fitness Testing in Physical Education – a Misdirected Effort in Promoting Healthy Lifestyles and Physical Activity?" *Physical Education and Sport Pedagogy* 14 (1): 89–108.

Carroll, B., and J. Loumidis. 2001. "Children's Perceived Competence and Enjoyment in Physical Education and Physical Activity Outside School." *European Physical Education Review* 7: 24–43.

Cecchini, J., C. González, Á Carmona, J. Arruza, A. Escartí, and G. Balagué. 2001. "The Influence of the Physical Education Teacher on Intrinsic Motivation, Self–Confidence, Anxiety, and pre-and Post-Competition Mood States." *European Journal of Sport Science* 1 (4): 1–11.

Chen, F. F. 2007. "Sensitivity of Goodness of fit Indexes to Lack of Measurement Invariance." *Structural Equation Modeling* 14 (3): 464–504.

Cheung, G. W., and R. B. Rensvold. 2009. "Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance." *Structural Equation Modeling* 9 (2): 233–255.

Corbin, C. B. 2002. "Physical Activity for Everyone: What Every Physical Educator Should Know About Promoting Lifelong Physical Activity." *Journal of Teaching in Physical Education* 21: 128–144.

Cox, A. E., and L. Williams. 2008. "Change in Physical Education Motivation and Physical Activity Behavior During Middle School." *Journal of Adolescent Health* 43 (5): 506–513.

Cox, A. E., S. Ullrich-French, J. Madonia, and K. Witty. 2011. "Social Physique Anxiety in Physical Education: Social Contextual Factors and Links to Motivation and Behavior." *Psychology of Sport and Exercise* 12 (5): 555–562.

Craft, L. L., T. M. Magyar, B. J. Becker, and D. L. Feltz. 2003. "The Relationship Between the Competitive State Anxiety Inventory–2 and Sport Performance: A Meta-Analysis." *Journal of Sport and Exercise Psychology* 25: 44–66.

Davidov, E., B. Meuleman, J. Cieciuch, P. Schmidt, and J. Billiet. 2014. "Measurement Equivalence in Cross-National Research." *Annual Review of Sociology* 40: 55–75.

Dishman, R. K., R. W. Motl, R. Saunders, G. Felton, D. S. Ward, M. Dowda, and R. R. Pate. 2005. "Enjoyment Mediates Effects of a School-Based Physical-Activity Intervention." *Medicine and Science in Sports and Exercise* 37 (3): 478–487.

Eccles, J. S., and R. W. Roeser. 2011. "Schools as Developmental Contexts During Adolescence." *Journal of Research on Adolescence* 21 (1): 225–241.

Finnish National Board of Education. 2014. *National Core Curriculum for Basic Education*. Helsinki: Finnish National Board of Education.

Goetz, T., N. C. Hall, A. C. Frenzel, and R. Pekrun. 2006. "A Hierarchical Conceptualization of Enjoyment in Students." *Learning and Instruction* 16 (4): 323–338.

Goudas, M., S. J. H. Biddle, and K. R. Fox. 1994. "Achievement Goal Orientations and Intrinsic Motivation in Physical Fitness Testing with Children." *Pediatric Exercise Science* 6: 159–167.

Gråsten, A. 2014. "Students' Physical Activity, Physical Education Enjoyment, and Motivational Determinants through a Three-Year School- Initiated Program." PhD diss., University of Jyväskylä

Hagger, M. S., N. L. D. Chatzisarantis, T. Culverhouse, and S. J. H. Biddle. 2003. "The "Processes by Which Perceived Autonomy Support in Physical Education Promotes Leisure-Time Physical Activity Intentions and Behavior: A Trans-Contextual Model."." *Journal of Educational Psychology* 95 (4): 784–795.

Harris, J., and L. Cale. 2006. "A Review of Children's Fitness Testing." *European Physical Education Review* 12 (2): 201–225.

Hashim, H., J. R. Grove, and P. Whipp. 2008. "Validating the Youth Sport Enjoyment Construct in High School Physical Education." *Research Quarterly for Exercise and Sport* 79 (2): 183–195.

Hembree, R. 1988. "Correlates, Causes, Effects and Treatment of Test Anxiety." *Review of Educational Research* 58 (1): 47–77.

Hopple, C., and G. Graham. 1995. "What Children Think, Feel, and Know About Physical Fitness Testing." *Journal of Teaching in Physical Education* 14: 408–417.

Hu, L. T., and P. M. Bentler. 1999. "Cutoff Criteria for fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus new Alternatives." *Structural Equation Modeling* 6: 1–55.

Huhtiniemi, M., A. Sääkslahti, A. Watt, and T. Jaakkola. 2019. "Associations among Basic Psychological Needs, Motivation and Enjoyment Within Finnish Physical Education Students." *Journal of Sports Science and Medicine* 18: 239–247.

Jaakkola, T. T., J. O. Liukkonen, T. Laakso, and Y. Ommundsen. 2008. "The Relationship Between Situational and Contextual Self-Determined Motivation and Physical Activity Intensity as Measured by Heart Rates During Ninth Grade Students' Physical Education Classes." *European Physical Education Review* 14 (1): 13–31.

Jaakkola, T., A. Sääkslahti, J. Liukkonen, and S. Iivonen. 2012. *"Peruskoululaisten Fyysisen Toimintakyvyn Seurantajärjestelmä" [Monitoring and Feedback System for Physical Functioning Capacity]*. Jyväskylä: Faculty of Sport and Health Sciences, University of Jyväskylä.

Jaakkola, T. T., A. Sääkslahti, S. Yli-Piipari, M. Manninen, A. Watt, and J. O. Liukkonen. 2013. "Student Motivation Associated With Fitness Testing in the Physical Education Context." *Journal of Teaching in Physical Education* 32: 270–286.

Kline, R. B. 2015. *Principles and Practice of Structural Equation Modeling*. New York: Guilford Press.

Kowal, J., and M. S. Fortier. 2000. "Testing Relationships from the Hierarchical Model of Intrinsic and Extrinsic Motivation Using Flow as a Motivational Consequence." *Research Quarterly for Exercise and Sport* 71 (2): 171–181.

Léger, L. A., D. Mercier, C. Gadoury, and J. Lambert. 1988. "The Multistage 20 m Shuttle run Test for Aerobic Fitness." *Journal of Sports Sciences* 6: 93–101.

Little, T. D. 2013. *Longitudinal Structural Equation Modeling*. New York: Guilfors Press.

Liukkonen, J., V. Barkoukis, A. Watt, and T. Jaakkola. 2010. "Motivational Climate and Students' Emotional Experiences and Effort in Physical Education." *Journal of Educational Research* 103 (5): 295–308.

Lodewyk, K. R., and A. Muir. 2017. "High School Females' Emotions, Self-Efficacy, and Attributions During Soccer and Fitness Testing in Physical Education." *The Physical Educator* 74 (2): 269–295.

Luke, M. D., and G. D. Sinclair. 1991. "Gender Differences in Adolescents' Attitudes Toward School Physical Education." *Journal of Teaching in Physical Education* 11 (1): 31–46.

Muthén, L. K., and B. O. Muthén. 2017. *Mplus User's Guide*. 8th ed. Los Angeles, CA: Muthén & Muthén.

Naughton, G. A., J. S. Carlson, and D. A. Greene. 2006. "A Challenge to Fitness Testing in Primary Schools." *Journal of Science and Medicine in Sport* 9 (1-2): 40–45.

Ntoumanis, N., and N. D. Myers. 2016. *An Introduction to Intermediate and Advanced Statistical Analyses for Sport and Exercise Scientists*. New York, NY: Wiley.

Ommundsen, Y., and S. E. Kvalø. 2007. "Autonomy–Mastery, Supportive or Performance Focused? Different Teacher Behaviours and Pupils' Outcomes in Physical Education." *Scandinavian Journal of Educational Research* 51 (4): 385–413.

Papaioannou, A., and O. Kouli. 1999. "The Effects of Task Structure, Perceived Motivational Climate and Goal Orientations on Students' Task Involvement and Anxiety." *Journal of Applied Sport Psychology* 11: 51–71.

Pustejovsky, J. E., and J. P. Spillane. 2009. "Question-order Effects in Social Network Name Generators." *Social Networks* 31: 221–229.

Rice, C. 2007. "Becoming 'The Fat Girl': Acquisition of an Unfit Identity." *Women Studies International Forum* 30 (2): 158–174.

Salin, K., and M. Huhtiniemi. 2018. "Physical Education in Finland - After Curriculum Reform 2016." In *Physical Education in Secondary School. Researches - Best Practices – Situation*, edited by S. Popovic, B. Antala, and J. Gardasevic, 329–334. Niksic: Federation Internationale D'Education Physique.

Sallis, J. F., T. L. McKenzie, M. W. Beets, A. Beighle, H. Erwin, and S. Lee. 2012. "Physical Education's Role in Public Health: Steps Forward and Backward Over 20 Years and HOPE for the Future." *Research Quarterly for Exercise and Sport* 83 (2): 125–135.

Satorra, A., and P. M. Bentler. 2001. "A Scaled Difference Chi-Square Test Statistic for Moment Structure Analysis." *Psychometrika* 66 (4): 507–514.

Scanlan, T. K., G. M. Chow, C. Sousa, L. A. Scanlan, and C. A. Knifsend. 2016. "The Development of the Sport Commitment Questionnaire-2 (English Version)." *Psychology of Sport and Exercise* 22: 233–246.

Scanlan, T. K., and J. P. Simons. 1992. "The Construct of Sport Enjoyment." In *Motivation in Sport and Exercise*, edited by G. Roberts, 199–215. Champaign, IL: Human Kinetics.

Schwarzer, R. 1986. "Self-Related Cognitions in Anxiety and Motivation: An Introduction." In *Self-Related Cognitions in Anxiety and Motivation*, edited by R. Schwarzer, 1–17. Hillsdale, NJ: Lawrence Erlbaum Associates.

SHAPE (Society of Health and Physical Educators). 2017. *Appropriate and Inappropriate Practices Related to Fitness Testing. Position Statement.* Reston, VA: SHAPE America, Society of Health and Physical Educators.

Silverman, S., X. D. Keating, and S. R. Phillips. 2008. "A Lasting Impression: A Pedagogical Perspective on Youth Fitness Testing." *Measurement in Physical Education and Exercise Science* 12: 146–166.

Simonton, K. L., K. Mercier, and A. C. Garn. 2019. "Do Fitness Test Performances Predict Students' Attitudes and Emotions Toward Physical Education?" *Physical Education and Sport Pedagogy* 24 (6): 549–564.

Soini, M. 2006. "Motivaatioilmaston yhteys aktiivisuuteen ja viihtymiseen liikuntatunneilla" [The associations between motivational climate, physical activity, and enjoyment in Finnish physical education]. PhD diss., University of Jyväskylä.

Telford, R. D., R. B. Cunningham, P. Waring, R. M. Telford, L. S. Olive, and W. P. Abhayaratna. 2013. "Physical Education and Blood Lipid Concentrations in Children: The LOOK Randomized Cluster Trial." *Plos One* 8 (10): 1–9. e76124.

Vallerand, R. J. 1997. "Toward a Hierarchical Model of Intrinsic and Extrinsic Motivation." In *Advances in Experimental Social Psychology*, edited by M. P. Zanna, 271–360. San Diego, CA: Academic Press.

Van De Schoot, R., P. Schmidt, A. De Beuckelaer, K. Lek, and M. Zondervan-Zwijnenburg. 2015. "Editorial: Measurement Invariance." *Frontiers in Psychology* 6: 1064.

Von der Embse, N., D. Jester, D. Roy, and J. Post. 2018. "Test Anxiety Effects, Predictors, and Correlates: A 30-Year Meta-Analytic Review." *Journal of Affective Disorders* 227: 483–493.

Wallhead, T. L., and J. Buckworth. 2004. "The Role of Physical Education in the Promotion of Youth Physical Activity." *Quest (grand Rapids, Mich)* 56 (3): 285–301.

Wigfield, A., S. L. Lutz, and A. L. Wagner. 2005. "Early Adolescents' Development Across the Middle School Years: Implications for School Counselors." *Professional School Counseling* 9 (2): 112–119.

Woodman, T., and L. Hardy. 2003. "The Relative Impact of Cognitive Anxiety and Self-Confidence upon Sport Perfornce: a Meta-Analysis." *Journal of Sports Sciences* 21 (6): 443–457.

Yli-Piipari, S., C. K. J. Wang, T. T. Jaakkola, and J. O. Liukkonen. 2012. "Examining the Growth Trajectories of Physical Education Students' Motivation, Enjoyment, and Physical Activity: A Person-Oriented Approach." *Journal of Applied Sport Psychology* 24 (4): 401–417.

Yli-Piipari, S., A. Watt, T. Jaakkola, J. Liukkonen, and J. E. Nurmi. 2009. "Relationships Between Physical Education Students' Motivational Profiles, Enjoyment, State Anxiety, and Self-Reported Physical Activity." *Journal of Sports Science & Medicine* 8: 327–336.

Zeidner, M. 1998. *Test Anxiety. The State of the Art*. New York: Plenum Press.