Students' attitudes towards and experiences of the Youth-fit health related fitness test battery.

Brendan T. O'Keeffe[a,b], Ciaran MacDonncha[a,b], and Alan E. Donnelly[a,b]

[a] *Department of Physical Education and Sport Sciences, University of Limerick;* [b] *Health Research Institute, University of Limerick.*

*Corresponding author: Brendan O'Keeffe (brendan.okeeffe@ul.ie)

This work was supported by the Irish Research Council under Grant [GOIPG/2017/789].

Students' attitudes towards and experiences of the Youth-fit health related fitness test battery.

The aim of this study was to examine secondary school students' attitudes towards and experiences of a student-centred health related fitness test battery. A total of 795 adolescents (403 boys, 50.7%; 392 girls, 49.3%) aged 13.2 years (±0.39) from 20 secondary schools in the Republic of Ireland participated in the study. Schools were stratified for gender, location and educational (dis)advantage. Students completed the test battery in small groups (n= ≤6) and each test item was administered by a trained senior student facilitator. Testing took place during physical education lessons. Test items included: body mass index; 20m shuttle run; back-saver sit and reach; hand-grip strength; standing long jump; isometric plank-hold; 90˚ push-up; 4x10m shuttle run; and blood pressure. Following participation in the test battery, students completed an instrument with valid scores for measuring attitudes towards fitness tests (Mercier and Silverman, 2014b). Students' experiences of each test item were also analysed. Overall, students had a positive attitude towards fitness testing (M= 3.9, ±0.59) on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Although both positive, the mean attitude score for boys (M=4.05, ±0.59) was significantly higher than girls (M=3.79, ±0.59; $p < 0.01$, *t*-test). Most students (n = 690, 86.8%) agreed or strongly agreed that the senior student facilitator made it easier for them to perform the tests. In conclusion, students had positive attitudes towards and experiences of the Youth-fit test battery. Physical education teachers should consider implementing a small group and senior student facilitated approach when administering fitness tests.

Keywords: adolescents, monitoring, health, physical fitness, student-centred

Introduction

Schools, and specifically physical education programmes, have been identified as the most suitable vehicle for the promotion of active and healthy lifestyles among young people (Pate et al., 2006). In recent years, the unique position of physical education as a key stakeholder in the promotion of healthy lifestyle behaviours among youth has been increasingly acknowledged by global (WHO, 2018) and national (Department of Health, 2016) policy makers. Rising obesity and inactivity levels among youth have stimulated a continued push for health-related fitness outcomes to be a focal point of physical education (Garn and Sun, 2009). In a worldwide survey of physical education programmes, health related physical fitness (HRPF) was ranked the number one most significant curriculum theme in school-based physical education (UNESCO, 2014). Fitness testing, a component of fitness education programmes, has been part of physical education for over half a century (Morrow et al., 2009), and has been reported as one of the most practical ways to teach the components of HRPF in physical education settings (Garn and Sun, 2009). Indeed, several states in the United States, and countries including Japan, China, Slovenia, Hungary and Finland have mandated monitoring physical fitness in school physical education programmes in the form of HRPF test batteries (O'Keeffe et al., 2020b).

Although fitness testing continues to be a component of most physical education programmes (Cooper et al., 2010), its role in physical education remains a divisive topic. Critics claim the prominence of fitness testing could be contributing to the growing performative culture that has enveloped education systems in recent years (Alfrey and Gard, 2014), and its inclusion in physical education programmes may well represent a misdirected effort at health promotion (Cale and Harris, 2009). In one of the few empirical studies of students' attitudes towards fitness testing, Hopple and Graham (1995)

suggested that students are often unclear about what they actually learn from the process. In contrast, proponents highlight the potential educational (Pate et al., 2013), motivational (Graser et al., 2011) and health (Ruiz et al., 2009) benefits of HRPF testing. Many scholars have cited the importance of teaching students how to self-assess their fitness levels as a core element of promoting lifelong physical activity (Castelli and Williams, 2007). Artero and colleagues (2011) suggested that fitness testing in school settings coordinated by qualified physical education professionals could represent a viable alternative to monitoring key indicators of health among youth in epidemiological studies. Despite these contrasting views, there is an emerging consensus that, if appropriately employed and incorporated as just one component of a broad and balanced programme, there is no reason why HRPF testing cannot make a valuable contribution to physical education in schools (Cale et al., 2014; Wiersma and Sherman, 2008).

There has been a wealth of research on HRPF test items that produce valid scores (Lubans et al., 2011; Ramírez-Vélez et al., 2015; Ruiz et al., 2009) in addition to pedagogically sound approaches to integrating fitness testing in school-based physical education programmes (Corbin et al., 2014; Harris and Cale, 2007; Liu, 2008; Wiersma and Sherman, 2008; Zhu et al., 2018). Recently, the ALPHA (Assessing Levels of Physical Activity) fitness test battery was developed to facilitate monitoring HRPF in a comparable way in the European Union (Ruiz et al., 2011). The ALPHA test battery was shown to produce valid and reliable scores, and was found to be safe for use in school settings when administered by a physical education teacher (Espana-Romero et al., 2010). Recommendations for integrating fitness testing in school-based physical education programmes in a pedagogically sound manner included: moving away from command-style test administration, to a more reciprocal, student-centred approach (Graser et al., 2011); using criterion rather than norm referenced health standards to promote self-

referenced comparison with attainable health standards (Mahar and Rowe, 2008); and allowing students an opportunity to familiarise themselves with tests prior to testing (Martin et al., 2010). However, a recent review of HRPF monitoring practices from a nationally representative sample of schools in the Republic of Ireland indicated that many teachers were not implementing these best practice recommendations, instead integrating tests in isolation and minimising learning opportunities for students (O'Keeffe et al., 2020b).

There is a dearth of research on student voice regarding their attitudes towards and experiences of fitness testing. Much of the research that is commonly cited when debating the role of fitness testing in physical education is based on a very small number of empirical research studies, some of which were conducted over quarter of a century ago (Cale and Harris, 2009; Hopple and Graham, 1995; Luke and Sinclair, 1991). Students' experiences of fitness tests using the aforementioned pedagogically sound recommendations have yet to be analysed in order to determine whether or not they actually work, and calls for more empirical work on the views of young people on fitness testing have mostly gone unanswered (Alfrey and Gard, 2019). Furthermore, given the prominence of testing in fitness education units (O'Keeffe et al., 2020b), what students think and feel about fitness testing and how these attitudes could affect participation in life-long physical activity are important to identify (Mercier and Silverman, 2014b). Vazou, Mischo and Ladwig (2019) similarly highlighted the scarcity of pragmatic experimental investigations examining the effects of specific changes to how physical education, and more specifically fitness testing, is delivered, on the quality of the experience that students derive.

In an effort to address this paucity of empirical research, Mercier and Silverman (2014b) developed an instrument to measure students' attitudes towards fitness testing.

The 18-item instrument, which was shown to produce valid and reliable scores with adolescent populations, is comprised of four sub-factors namely: cognitive; affect-enjoyment; affect-feelings; and affect-teacher (Mercier and Silverman 2014a). The instrument was informed by Ajzen's (1991) Theory of Planned Behaviour, which posits that attitudes are a significant predictor of human behaviour, in addition to subjective norms and perceived behavioural control. Researching attitudes is important as decisions, such as whether to remain physically active, can be strongly influenced by attitudes (Solmon, 2003). In the only large scale study to use the instrument to date, Mercier and Silverman (2014a) found that students had a slightly positive attitude toward fitness testing, with boys reporting significantly higher mean values in comparison to girls. However, this instrument has not been used to examine students' attitudes towards fitness testing outside of the United States. In addition, to the authors' knowledge, no study has examined students' experiences of multiple commonly administered HRPF test items when delivered in the form of a test battery in a physical education context. Therefore, the aim of the current study was to examine students' attitudes towards and experiences of a student-centred health related fitness test battery in secondary school-based physical education programmes.

Methods

The methodological design of this study included four steps: 1) instrumentation; 2) school and participant recruitment; 3) Youth-fit test battery administration and evaluation; 4) data analysis. Ethics committee approval was granted by the Research Ethics Committee of the Faculty of Education and Health Sciences at the University of Limerick (EHS_2017_02_12).

Instrumentation

Data were collected using the Students' Attitudes towards Fitness Testing instrument, developed by Mercier and Silverman (2014b). This is an 18-item instrument, made up of four sub-factors: cognitive (6 items), affect-enjoyment (3 items), affect-feelings (4 items) and affect-teacher (5 items). Scores from this instrument were shown to be valid and reliable for measuring adolescents' attitudes towards fitness testing. The development of this instrument has been described comprehensively elsewhere (Mercier and Silverman, 2014b). Questions specifically in relation to students' experiences of the Youth-fit test battery were included in the evaluation in addition to the 18-item instrument used to examine attitudes towards fitness tests. Students were asked to rate their experiences of completing each test item on a scale of 1 (very poor) to 5 (very good). Students also indicated if they preferred a self/peer, teacher or external expert approach to administering the test battery, and if they shared their HRPF results with a parent/guardian. Finally, students were asked two open ended questions which required them to identify the most and least enjoyable component of the test battery.

A confirmatory factor analysis (CFA) was used to examine data fit to the model proposed by Mercier and Silverman (2014b) for the current study sample using SPSS Amos (v26, Chicago IL). The CFA confirmed an overall good fit of the data to the four-

factor model; all indicator variables loaded significantly (p <.001) on the associated latent factor. Model fit indices including the comparative fit index, the Tucker-Lewis Index and root mean square error of approximation, were .897, .877 and .085, respectively, indicating a good fit of the data to the model. Cronbach's alpha internal consistency coefficients were also determined and represented good to excellent levels of reliability for each sub-factor and for the overall model. The alpha reliability coefficient for the entire model was .892, and the four factors and their reliability scores were cognitive (α .887), affective-enjoyment (α .808), affective-feelings (α .834), and affective-teacher (α .732). Furthermore, a Cronbach's alpha measure of .851 was established for the 10-item test experience scale developed specifically for this study, representing excellent reliability.

School and participant recruitment

A randomised sample of 20 schools, stratified for gender (boys, girls and mixed-gender), location (categorised by population density: urban, the cities of Cork and Limerick; rural, all other areas of the mid and south west of the Republic of Ireland) and educational (dis)advantage (designated disadvantaged and non-disadvantaged schools), classified by the Department of Education and Skills, Government of Ireland (2017), participated in the study. If a school in the initial sample declined to participate, a replacement list of schools for each stratum was generated to provide an alternative school with the same demographics. Due to the geographical spread of schools, and the need to visit each school individually, 20 schools was considered to be the maximum sample size achievable from a logistical viewpoint, and the minimum required to ensure a sufficient number of schools in each of the chosen strata. Approval from the principal and cooperating physical education teacher in each school was granted following an initial

email and telephone conversation. Written informed consent was obtained from the parents/guardians of the students, and the students themselves. The Republic of Ireland education system comprises three levels including, primary, post-primary (secondary) and third level. Post primary or secondary education is made up of junior cycle (years one to three, ages 13 to 15), a transition year (year four, age 16) and senior cycle (years five and six, ages 17 to 18). This study focused specifically on students in year one of secondary school education (ages 13 to 14) and was open to all students in the selected year group in each participating school.

Youth-fit test battery administration and evaluation

Cooperating physical education teachers in each school selected eight senior students (years four to six, ages 16 to 18) as test facilitators. A detailed standard operating procedure was designed for and read by senior student facilitators and cooperating teachers, who also participated in a three-hour workshop delivered by the lead author one week in advance of testing. Administration protocols for each test item have been described elsewhere (O'Keeffe et al., 2020a). One week after administering the test battery, participating students completed the evaluation survey during physical education class time. For convenience and wide distribution, an online questionnaire, via SurveyMonkey (Palo Alto, CA) cloud-based software, was utilised to administer the evaluation survey. All student participants received an email that outlined the purpose of the survey, details regarding the time commitment and confidentiality, and a web link to complete the survey. Participants were informed that they could exit the survey at any time without implication. Cooperating physical education teachers clarified any questions students had on specific items in the survey.

Data analysis

Complete responses (*N* =795) were extracted from SurveyMonkey and transferred to SPSS (version 25; IBM Corp. Chicago, IL) for analysis. The research team defined an incomplete response as missing one or more items from the attitude instrument. Incomplete responses (*n* =74) were excluded from all analyses. A visual inspection of histograms for key outcome variables showed that data were normally distributed, with skewness of ≤ -1.3 (SE = .087) and kurtosis of ≤ 1.9 (SE = .173). Descriptive statistics including means (M), standard deviations (±) and 95% confidence intervals were determined for gender and school type for factor and overall attitude scores. Independent samples *t*-tests were used to compare differences between boys and girls, and school level socio-economic status. Effect size was determined by calculating the mean difference between the two groups and dividing the result by the pooled standard deviation (Cohen's *d*).

A factorial multivariate analysis of variance (MANOVA) using gender and school type as independent variables and the four factor variables as dependent variables was performed. The absence of multicollinearity was confirmed by examining correlations among the four sub-factor dependent variables. The dependent variables were moderately related, ranging from r =.27 to r =.58, and thus did not exceed the .80 threshold (Dormann et al., 2013). A separate ANOVA was conducted for each factor. Effect sizes for ANOVAs were calculated using the η^2 (Eta squared) method, by dividing the treatment sum of squares by the total sum of squares. A one sample *t*-test, was used to compare total and factor mean values of students in this study with age-matched values reported in a study that used the same instrument to measure students' attitudes towards fitness tests in the United States (Mercier and Silverman, 2014a). The rationale for this comparison was to determine if there was a significant difference in students' responses

189 to contrasting approaches of administering fitness test batteries in school contexts.
190 Correction for multiple comparisons was via the Bonferroni correction (Abdi, 2007).

191 Responses to the two open-ended questions regarding the most and least enjoyable
192 aspect of the test battery were reviewed and organised thematically in line with the
193 guidelines set out by Taylor-Powell and Renner (2003) for analysing qualitative data.
194 This involved identifying themes and patterns from the responses. Once the key themes
195 had been established and agreed upon by each author, responses were arranged into
196 coherent categories and frequencies of responses within each category were calculated.

Results

Overall, students had a positive attitude towards fitness testing (M= 3.92, ±0.59) on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The cognitive (perceived usefulness) factor had the highest mean score (M=4.19, ±0.66), while the affect-feelings factor produced the lowest mean score (M=3.58, ±1.05). Although both positive, boys (M=4.05, ±0.55) had significantly higher overall attitude scores in comparison to girls (M=3.79 ±0.59) (t (793) = 6.34, p = 0.001, Cohen's d = .44). Boys produced significantly higher scores across all factors in comparison to girls, with the exception of affect-teacher, in which there was no significant difference. The cognitive factor produced the highest mean scores for boys and girls; however, the lowest mean scores differed between genders, affect-teacher for boys and affect-feelings among girls, as illustrated in **Table 1**.

Table 1. A comparison of overall and sub-factor descriptive statistics of boys' and girls' attitudes towards fitness testing.

Variable	**Total Mean (±)** (N=795)	**Boys Mean (±)** (N=403)	**Girls Mean (±)** (N=392)	**95% CI of difference (lower)**	**95% CI of difference (higher)**	***p* value**	***Effect size*** ***(Cohen's d)***
Overall	3.92 (.59)	4.05 (.55)	3.79 (.59)	0.18	0.34	<0.01*	.44
Cognitive	4.19 (.66)	4.27 (.66)	4.12 (.64)	0.05	0.24	<0.01*	.22
Affect Enjoyment	3.90 (.83)	4.08 (.77)	3.71 (.86)	0.26	0.49	<0.01*	0.45
Affect Teacher	3.82 (.68)	3.86 (.67)	3.78 (.69)	0.01	0.18	NS	0.12
Affect Feelings	3.58 (1.05)	3.88 (.95)	3.27 (1.05)	0.47	0.75	<0.01*	0.58

Abbreviations: CI, confidence interval. Data are shown as means with standard deviation in brackets. * Significant at Bonferroni adjusted p value < 0.01.

In terms of differences by selected demographics, total and factor attitude scores for students in designated disadvantaged and non-disadvantaged schools did not differ significantly, with the exception of the cognitive factor which produced a significantly higher mean score for students in non-disadvantaged schools (M=4.22, ±0.65) in

comparison to those in designated disadvantaged schools (M=4.08, ±0.65) (t (793) = 2.58, p = 0.01, Bonferroni correction). Furthermore, total and factor attitude scores did not differ significantly between students from urban or rural schools.

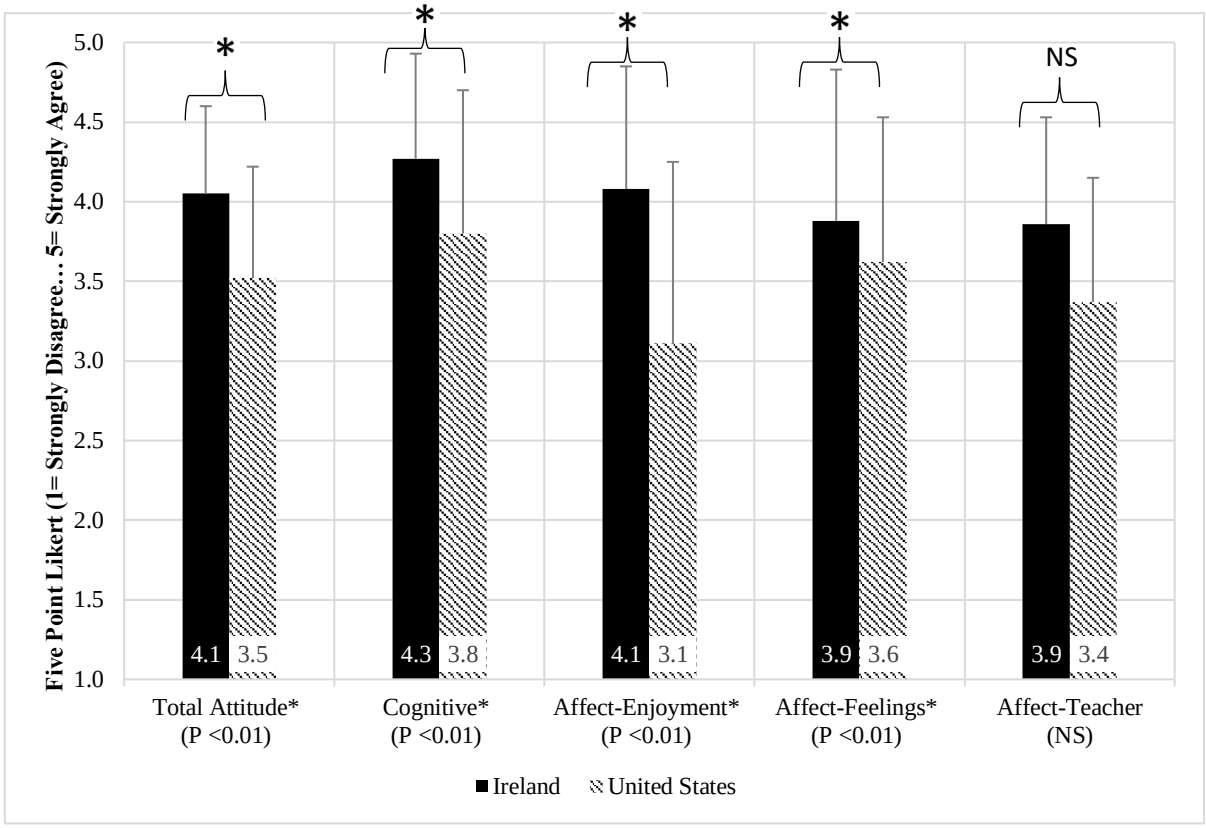
A MANOVA with follow-up, using gender and school type as independent variables and the four factor variables as dependent variables indicated that there was a significant difference between participants in boys, girls and mixed-gender schools when considered through each of the four factors combined, Wilk's Λ = .970, F(8, 1578) = 3.07, p =.002, partial $\eta 2$ = 0.15. A separate ANOVA was conducted for each factor using a Bonferroni adjusted p value of .01. Post-hoc comparisons using the Tukey HSD test indicated that the total attitude mean scores for boys' schools and mixed-gender schools were significantly higher than girls' schools, specifically within the cognitive, affect-enjoyment and affect-feeling factors. However, despite reaching statistical significance, the actual differences in mean values between school types was quite small, as evidenced by the effect size values in **Table 2**. Although relatively small, the largest differences were produced in the affect feelings factor F(2, 792) = 10.00, p = .001, $\eta 2$ = 0.03. A significant difference was not found between school types in the affect teacher domain, F(2,792) = 2.91, p = .06, $\eta 2$ = 0.001.

Table 2. Overall and sub-factor mean (±) scores for students' attitudes towards fitness testing by school type.

Dependent Variable	School Type	Mean (±)	Lower (95% CI)	Upper (95% CI)	*p* value	*Effect size (η2)*
Cognitive	Boys (n = 100)	4.27 (0.51)	4.14	4.40	<0.01*	.01
	Girls (n = 101)	4.02 (0.75)	3.89	4.15		
	Mixed (n = 594)	4.21 (0.66)	4.16	4.26		
Affect Enjoyment	Boys	3.98 (0.82)	3.82	4.14	<0.01*	.02
	Girls	3.63 (0.90)	3.46	3.79		
	Mixed	3.93 (0.82)	3.87	4.00		
Affect Teacher	Boys	3.85 (0.66)	3.72	3.99	NS	.00
	Girls	3.67 (0.77)	3.54	3.80		
	Mixed	3.84 (0.66)	3.79	3.90		
Affect Feelings	Boys	3.81 (0.95)	3.61	4.02	<0.01*	.03
	Girls	3.19 (1.1)	2.99	3.39		
	Mixed	3.61 (1.0)	3.53	3.69		
Overall	Boys	4.01 (0.48)	3.92	4.11	<0.01*	.02
	Girls	3.76 (0.67)	3.56	3.83		
	Mixed	3.95 (0.58)	3.90	3.99		

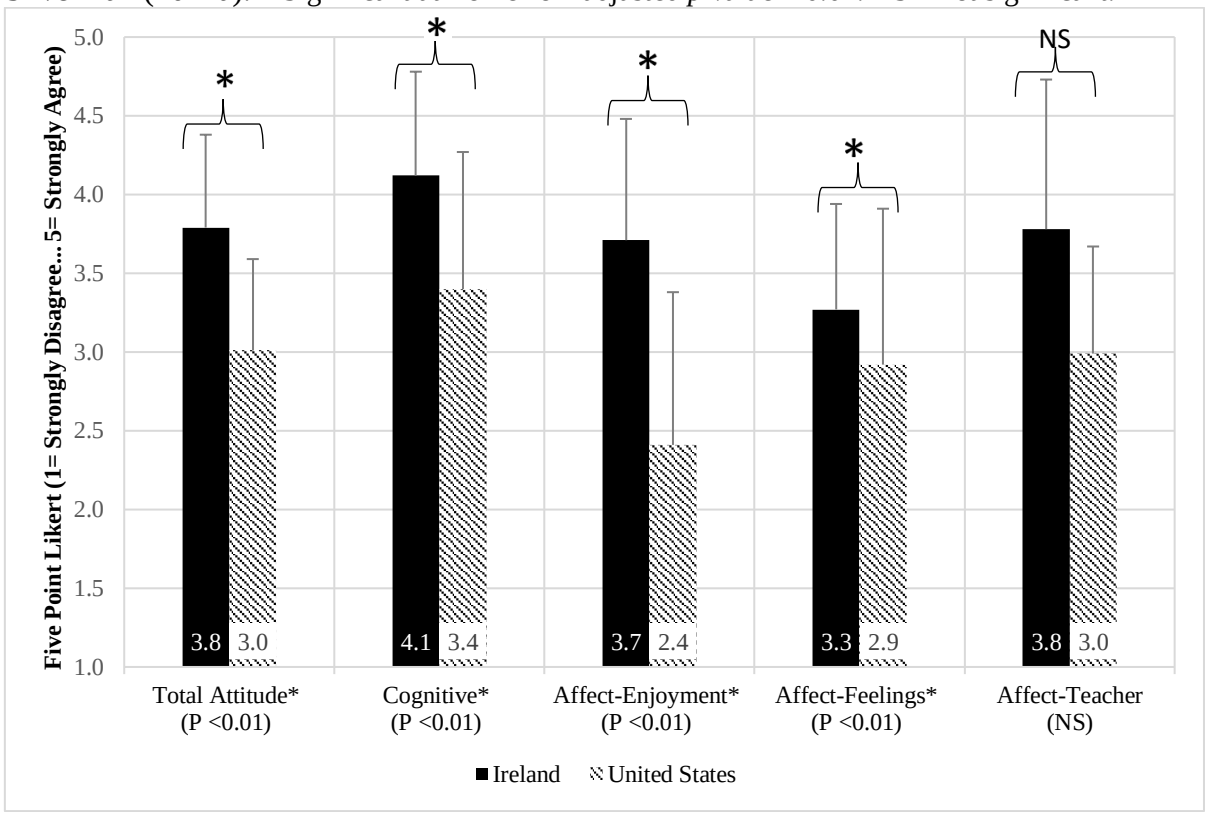
Abbreviations: CI, confidence interval. Data are shown as means with standard deviation in brackets. * Significant at Bonferroni adjusted *p* value < 0.01.

When compared with single value age-matched mean scores from United States, as reported in Mercier and Silverman (2014a), boys (**Figure 1**) and girls (**Figure 2**) in the Youth-fit study had a significantly higher overall mean attitude scores across (t (793) = 2.58, p = 0.01). Participants in this study scored significantly higher across three of the four sub-factors, with the largest difference recorded in the affect-enjoyment factor for both boys (t (403) = 25.8, p < .001) and girls (t (392) =14.0, p < .001).



248

249 **Figure 1.** A comparison of overall and sub-factor attitude scores between boys in the
250 Youth-fit study (Ireland) and single value age-matched mean scores from Mercier and
251 Silverman (2014a). *Significant at Bonferroni adjusted *p* value < 0.01. NS = not significant.



252

Figure 2. A comparison of overall and sub-factor attitude scores between girls in the Youth-fit study (Ireland) and single value age-matched mean scores from Mercier and Silverman (2014a). *Significant at Bonferroni adjusted *p* value < 0.01. NS = not significant.

Responses to survey items examining students' experiences of the Youth-fit test battery specifically were encouraging. Overall, 78% (n=618) of students agreed or strongly agreed that completing the Youth-fit test battery was a worthwhile experience, with a further 16% undecided. The vast majority of students (n = 636, 81.3%) agreed or strongly agreed that they would like to track their HRPF while in secondary school. When asked to rank who they would like to administer fitness tests from most preferred to least preferred, students (n = 413, 52.8%) indicated that they would be in favour of the student-centred format used in the Youth-fit test battery, in comparison to an external expert (27.0%) or their teacher (20.2%) recording test scores. In addition, 86.8% (n = 690) of students agreed or strongly agreed that the senior student facilitator made it easier for them to perform each test item. Students reported fair to good experiences of each test item on a five-point Likert scale ranging from very poor to very good (M= 3.78 ±1.0). The 90° push-up for girls (M= 3.24 ±1.27) and sit and reach for boys (M= 3.52 ±1.16) had the lowest mean scores, while the 4x10m shuttle run had the highest mean scores for both groups (Girls, M=3.91 ±0.90; Boys , M= 4.19 ±0.89). Overall, boys (M=3.45 ±0.62) reported significantly more positive test experiences in comparison to girls (M=3.33 ±0.58) on all items combined ($t(766) = 2.73$, $p = .007$). It was also interesting to note that students who indicated they shared their test results with a parent/guardian had a significantly higher mean score on the attitude towards fitness tests instrument (M=4.02 ±0.49) in comparison to those who did not (M=3.69 ±0.62) ($t(218) = 4.4$, $p = 0.001$).

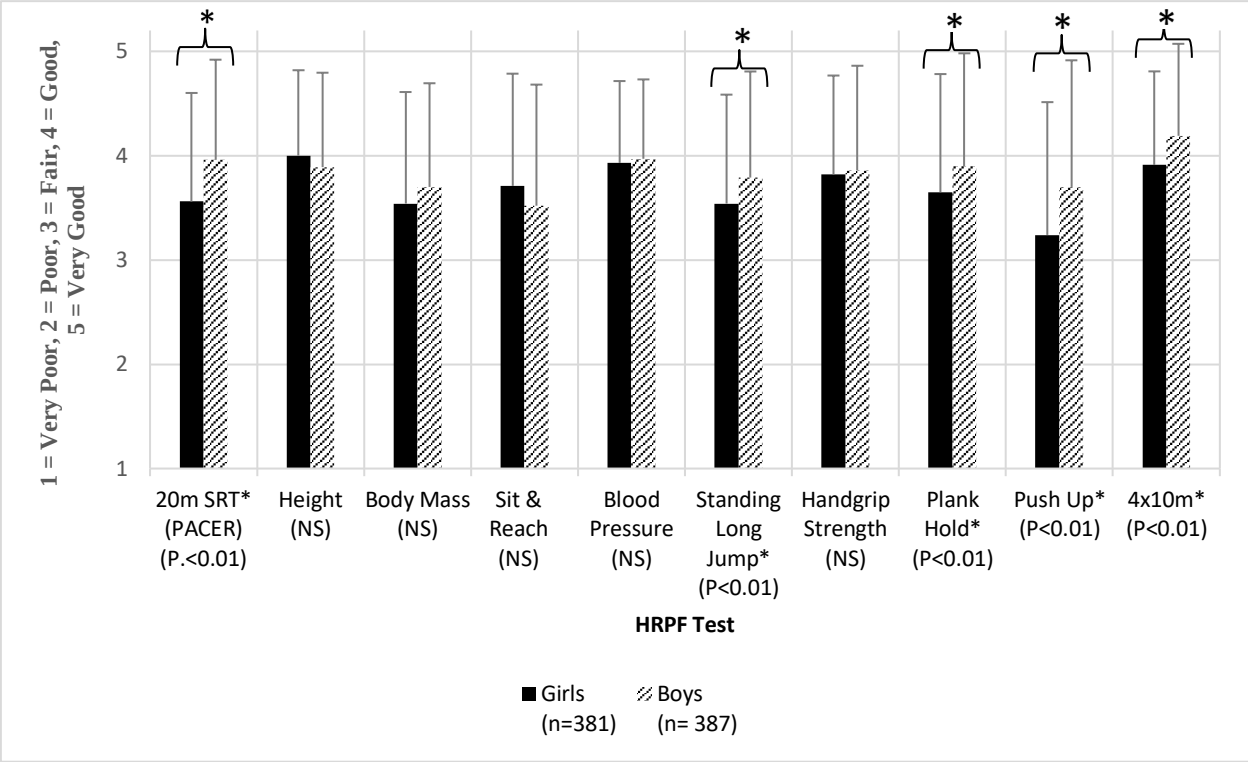


Figure 3. A comparison of boys' and girls' experiences of each Youth-fit test battery item. *Note:* *Denotes a significant difference, $p < 0.01$.

The final part of the evaluation survey was comprised of two open ended style questions in which students were requested to identify the most and least enjoyable part of the test battery. Participating with friends and having fun were the most commonly cited enjoyable aspects of the test battery (n=196). For example, one student stated, (I enjoyed) "*doing it with a small group of my closest friends and motivating each other to get a good score*" (male, boys school). Completing the muscular endurance test items was cited by 210 participants as the least enjoyable component of the test battery. Although not in the top five most frequently cited least enjoyable components of the test battery, 83 students noted having their height/weight measured. One student stated, "*if someone is a bit heavier than another then they could get bullied for their weight*" (male, mixed school).

Discussion

The aim of this study was to examine secondary school students' attitudes towards and experiences of a student-centred health related fitness test battery. Much of the research on students' perceptions of fitness testing to date have been generated from relatively small sample sizes (Davis et al., 2018; Garn and Sun, 2009; Graser et al., 2011), and participants in these studies often experienced different test items and administration protocols (Mercier and Silverman, 2014a). This investigation, the first outside of the United States to use the students' attitudes towards fitness testing instrument (Mercier and Silverman, 2014b), included a stratified sample of students from a randomised sample of 20 secondary schools. In summary, students had a positive attitude towards fitness testing, boys had a significantly more positive attitude than girls, and students in this study produced significantly higher mean scores when compared to age-matched scores of students from the United States whose attitudes towards fitness testing were measured using the same instrument.

Students had a positive attitude towards fitness tests, with the cognitive, or perceived usefulness, factor scoring highest across all demographic groupings, indicating that they perceived fitness testing to be a useful component of their physical education programme. This corroborates the findings of Mercier and Silverman (2014a), in which the same instrument was administered to 1199 students from 9^{th} to 12^{th} grade in the United States. Unlike Mercier and Silverman's study, in which affect-enjoyment was the factor with the lowest mean score, the affect-feelings factor had the lowest mean score among students in this study, indicating that students may have been nervous about performing the tests. However, relatively large standard deviations for each factor reported by Mercier and Silverman, likely composed of students with highly positive or highly

316 negative attitudes, were not as extreme in this investigation, indicating less variation in
317 students' attitudes. Indeed, students in this study had significantly higher mean attitude
318 scores when compared to age-matched data from the United States in three of the four
319 factor variables (Mercier and Silverman, 2014a), namely, cognitive, affect-feelings and
320 affect-enjoyment.

321 An interesting finding to emerge from the current study was the disparity between
322 boys' and girls' attitudes towards fitness testing. Boys reported significantly more
323 positive attitudes across all factors in comparison to girls, with the exception of affect-
324 teacher which did not differ significantly. Furthermore, mean scores for students in boys'
325 schools and mixed-gender schools were significantly higher than girls' schools. An
326 exploratory study which investigated the factors that influence high school girls'
327 enrolment in elective physical education (Davis et al., 2018) indicated that students
328 (n=17) acknowledged the importance of HRPF, but desired less of a focus on fitness
329 testing. Similarly, Zhu, Chen and Parrott (2014) found that boys reported significantly
330 higher situational interest in the PACER (progressive aerobic cardiovascular endurance
331 run) test in comparison to girls. Engaging girls in physical education has been reported
332 as a challenge for teachers across most aspects of a curriculum (Enright and O'Sullivan,
333 2010), particularly so when it comes to fitness testing (Davis et al., 2018). However,
334 although girls reported significantly lower attitudes in comparison to boys, their mean
335 scores were positive and significantly higher than those reported by Mercier and
336 Silverman (2014a). This suggests that physical education teachers should consider
337 adopting the student-centred test administration protocol used in this study. However,
338 despite the positive attitudes reported, further research is needed on how best to integrate
339 fitness tests to ensure that girls in particular are comfortable participating and motivated
340 to try their best.

Although fitness testing is highly prevalent in physical education programmes internationally, to the authors' knowledge, this is the first study of its kind to quantitatively analyse students' experiences of multiple HRPF test items. Students reported fair to good experiences of each test, with the lowest mean scores recorded in the 90° push-up for girls and sit and reach test for boys. It has been suggested that the maximal and physically challenging nature of aerobic tests, such as the PACER, may lead to more negative motivation among students, and the appropriateness of such tests in a physical education context has been questioned (Cale et al., 2014; Ladwig et al., 2018; Wrench and Garrett, 2008). Interestingly, students in this study reported fair to good experiences of the PACER test, and it was frequently highlighted as the most enjoyable aspect of the test battery in the open-ended part of the survey. Simonton and colleagues (2019) recently reported that PACER performance predicted lower reports of future anger toward physical education for both girls and boys.

It should also be noted that the body mass test item had the joint second lowest mean score among students, and having body mass and height recorded was cited by 83 students as the least enjoyable aspect of the test battery. While some scholars have highlighted the benefits of systematic monitoring of anthropometric measures in school settings (Thompson et al., 2019), physical education teachers need to be mindful that body image concerns and elevated levels of anxiety appear to undergird the influence of self-efficacy in fitness test performance, particularly so in females (Lodewyk and Sullivan, 2016). Lodewyk and Sullivan (2016) provide some suggestions to assist physical education teachers in structuring fitness education units to better minimise this vulnerability. Recommendations included, reducing social and normative comparisons, making accommodations for attire and providing the option of gender-segregated health-related fitness units in coeducational settings. However, the recommendation to offer

366 gender-segregated HRPF units could be questioned based on the findings of this study
367 which indicated that girls in coeducational or mixed-gender schools had significantly
368 more positive attitudes towards fitness testing than participants in girls' schools.
369 Providing students in coeducational settings with the opportunity to complete fitness tests
370 battery in small groups of their closest peers regardless of gender could potentially
371 enhance students', and girls' in particular, motivation to participate.

372 Much of the existing research regarding students' attitudes towards fitness
373 education, and physical education more broadly, has highlighted the importance of
374 enjoyment (Garn and Sun, 2009; Prochaska et al., 2003). Participating with friends and
375 having fun was the most commonly cited enjoyable aspect of the HRPF test battery by
376 students in this study. Similar studies that have investigated experiences of fitness tests
377 to date indicate students enjoy a peer-assessed testing format in small groups (Mercier
378 and Silverman, 2014a; Phillips et al., 2017). The student-centred approach affords
379 participants the opportunity to develop a sense of personal control and fitness autonomy
380 (Biddle and Fox, 1998). Indeed, Prusak and Vincent (2005) noted that students tested in
381 an environment that supports their autonomy will be more motivated to participate and
382 strive for self-improvement. Research by O'Keeffe et al. (2020a) demonstrated that,
383 following a period of familiarisation with test items, student administered fitness tests in
384 physical education lessons were as reliable as those taken by experienced research
385 assistants. The vast majority of students in the current study agreed that the senior student
386 facilitator at each test station made it easier for them to complete each test. The student-
387 centred approach also offers the physical education teachers the opportunity to move
388 throughout the learning space, and provide individualised feedback where appropriate,
389 without being restricted to delivering a single test item. Therefore, both the process of

390 engaging in the test battery and the validity of the scores obtained can be simultaneously
391 enhanced.

392 Students who shared their fitness test results with a parent/guardian had
393 significantly more positive attitudes towards fitness testing in comparison to those who
394 did not share their results. A recent review of HRPF monitoring practices involving a
395 nationally representative sample of schools in the Republic of Ireland revealed that less
396 than one third of physical education teachers shared HRPF test results with their students'
397 parents/guardians (O'Keeffe et al., 2020b). Mercier et al. (2016) also reported that less
398 than 30% of teachers surveyed in their study sent HRPF test results home. The importance
399 of integrating parents/guardians in a child's education has been well established (Jeynes,
400 2007), and sharing fitness test results could represent a useful avenue to keep
401 parents/guardians informed of their child's HRPF levels. Furthermore, many school-
402 based fitness test batteries, including Fitnessgram (Meredith and Welk, 2010), have been
403 updated to include a greater focus on physical activity promotion in addition to physical
404 fitness, further emphasising the opportunity to integrate fitness testing as part of a broader
405 HRPF education unit and more broadly to promote lifelong physical activity beyond the
406 school context.

407 Caution should be applied when interpreting the results of this study. Participants
408 in the current study were in year one of secondary school education and research
409 consistently indicates that students' attitudes towards physical education decline as they
410 get older (Silverman, 2017). Furthermore, comparisons with data that used the same
411 measurement instrument in the United States (Mercier and Silverman, 2014a) were
412 generated from age-matched fixed means as opposed to the original dataset. Future
413 research should analyse how attitudes towards fitness testing change as students progress
414 through secondary school. In addition, although students in each of the 20 schools

involved in the study were represented, response rates within schools varied between 60 and 100 percent, which could have resulted in some response bias. Finally, this survey was based on a single data source; therefore, data were not verified a second time, potentially resulting in a lack of depth in the interpretation of responses for which qualitative methods may have provided further insights.

Conclusion

Analysing students' attitudes and experiences is a critical step in developing evidence-based pedagogical approaches. Overall, students had a positive attitude towards fitness testing and participants clearly perceived fitness testing to be a useful component of a fitness education unit. This study illustrates the potential of a student-centred approach to administering fitness tests in a physical education context. Teachers should strongly consider educating senior students to facilitate in both the set-up and administration of a fitness test battery as a mechanism to enhance participants' understanding of each test item while also improving the accuracy of the measure obtained. Further research is needed to confirm if the positive responses to the student-centred approach presented in the current study maintain as students progress through secondary school.

Acknowledgement

The study received support funding from the Government of Ireland Postgraduate Scholarship Scheme.

References

Abdi H (2007) Bonferroni and Šidák corrections for multiple comparisons. *Encyclopedia of measurement and statistics* 3: 103-107.

Ajzen I (1991) The theory of planned behavior. *Organizational behavior and human decision processes* 50(2): 179-211.

Alfrey L and Gard M (2014) A crack where the light gets in: a study of Health and Physical Education teachers' perspectives on fitness testing as a context for learning about health. *Asia-Pacific Journal of Health, Sport and Physical Education* 5(1): 3-18.

Alfrey L and Gard M (2019) Figuring out the prevalence of fitness testing in physical education: A figurational analysis. *European Physical Education Review* 25(1): 187-202.

Artero EG, Espana-Romero V, Castro-Pinero J, et al. (2011) Reliability of field-based fitness tests in youth. *International Journal of Sports Medicine* 32(3): 159-169.

Biddle S and Fox K (1998) Motivation for physical activity and weight management. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity* 22: S39-47.

Cale L and Harris J (2009) Fitness testing in physical education – a misdirected effort in promoting healthy lifestyles and physical activity? *Physical Education & Sport Pedagogy* 14(1): 89-108.

Cale L, Harris J and Chen MH (2014) Monitoring health, activity and fitness in physical education: its current and future state of health. *Sport, Education & Society* 19(4): 376-397.

Castelli D and Williams L (2007) Health-Related Fitness and Physical Education Teachers' Content Knowledge. *Journal of Teaching in Physical Education* 26(1): 3-19.

Cooper KH, Everett D, Meredith MD, et al. (2010) Preface: Texas Statewide Assessment of Youth Fitness. *Research Quarterly for Exercise & Sport* 81: ii-iv.

Corbin CB, Welk GJ, Richardson C, et al. (2014) Youth Physical Fitness: Ten Key Concepts. *Journal of Physical Education, Recreation & Dance* 85(2): 24-31.

Davis S, Zhu X and Haegele J (2018) Factors influencing high school girls' enrolment in elective physical education: an exploratory qualitative inquiry. *Curriculum Studies in Health and Physical Education*. 1-14.

Department of Education and Skills (2017) Delivering Equality of Opportunity in Schools: DEIS Plan 2017. Dublin: Government of Ireland.

Department of Health (2016) Healthy Ireland: Get Ireland Active. National Physical Activity Plan for Ireland. Dublin: Government of Ireland.

Dormann CF, Elith J, Bacher S, et al. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36(1): 27-46.

Enright E and O'Sullivan M (2010) 'Can I do it in my pyjamas?'Negotiating a physical education curriculum with teenage girls. *European Physical Education Review* 16(3): 203-222.

Espana-Romero V, Artero EG, Jimenez-Pavon D, et al. (2010) Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA Study. *Int J Sports Med* 31(7): 490-497.

Garn A and Sun H (2009) Approach-avoidance motivational profiles in early adolescents to the PACER fitness test. *Journal of Teaching in Physical Education* 28(4): 400-421.
Graser SV, Sampson BB, Pennington TR, et al. (2011) Children's Perceptions of Fitness Self-Testing, the Purpose of Fitness Testing, and Personal Health. *Physical Educator* 68(4): 175-187.
Harris J and Cale L (2007) Children's fitness testing: A feasibility study. *Health Education Journal* 66(2): 153-172.
Hopple C and Graham G (1995) What children think, feel, and know about physical fitness testing. *Journal of Teaching in Physical Education* 14(4): 408-417.
Jeynes WH (2007) The Relationship Between Parental Involvement and Urban Secondary School Student Academic Achievement:A Meta-Analysis. *Urban Education* 42(1): 82-110.
Ladwig MA, Vazou S and Ekkekakis P (2018) "My best memory is when I was done with it": PE memories are associated with adult sedentary behavior. *Translational Journal of the American College of Sports Medicine* 3(16): 119-129.
Liu Y (2008) Youth Fitness Testing: If the "Horse" is Not Dead, What Should We Do? *Measurement in Physical Education and Exercise Science* 12(3): 123-125.
Lodewyk KR and Sullivan P (2016) Associations between anxiety, self-efficacy, and outcomes by gender and body size dissatisfaction during fitness in high school physical education. *Physical Education and Sport Pedagogy* 21(6): 603-615.
Lubans DR, Morgan P, Callister R, et al. (2011) Test-retest reliability of a battery of field-based health-related fitness measures for adolescents. *J Sports Sci* 29(7): 685-693.
Luke MD and Sinclair GD (1991) Gender differences in adolescents' attitudes toward school physical education. *Journal of Teaching in Physical Education* 11(1): 31-46.
Mahar MT and Rowe DA (2008) Practical Guidelines for Valid and Reliable Youth Fitness Testing. *Measurement in Physical Education and Exercise Science* 12(3): 126-145.
Martin SB, Ede A, Morrow JR, et al. (2010) Statewide Physical Fitness Testing. *Research Quarterly for Exercise and Sport* 81(sup3): S31-S41.
Mercier K, Phillips S and Silverman S (2016) High School Physical Education Teachers' Attitudes and use of Fitness Tests. *High School Journal* 99(2): 179-190.
Mercier K and Silverman S (2014a) High school students' attitudes toward fitness testing. *Journal of Teaching in Physical Education* 33(2): 269-281.
Mercier K and Silverman S (2014b) Validation of an Instrument to Measure High School Students' Attitudes Toward Fitness Testing. *Research Quarterly for Exercise & Sport* 85(1): 81-89.
Meredith MD and Welk G (2010) *Fitnessgram and Activitygram Test Administration Manual-Updated 4th Edition.* Dallas: Human Kinetics.
Morrow JR, Zhu W, Franks DB, et al. (2009) 1958–2008: 50 years of youth fitness tests in the United States. *Research Quarterly for Exercise and Sport* 80(1): 1-11.
O'Keeffe BT, Donnelly AE and MacDonncha C (2020a) Test-retest reliability of student administered health-related fitness tests in school settings. *Pediatric Exercise Science* 32(1): 48-57.

O'Keeffe BT, MacDonncha C, Ng K, et al. (2020b) Health-Related Fitness Monitoring Practices in Secondary School-Based Physical Education Programmes. *Journal of Teaching in Physical Education* 39(1): 59-68.

Pate RR, Davis MG, Robinson TN, et al. (2006) Promoting physical activity in children and youth: a leadership role for schools: a scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Physical Activity Committee) in collaboration with the Councils on Cardiovascular Disease in the Young and Cardiovascular Nursing. *Circulation* 114(11): 1214-1224.

Pate RR, Welk GJ and McIver KL (2013) Large-Scale Youth Physical Fitness Testing in the United States: A 25-Year Retrospective Review. *Pediatric Exercise Science* 25(4): 515-523.

Phillips SR, Marttinen R and Mercier K (2017) Fitness Assessment: Recommendations for an Enjoyable Student Experience. *Strategies* 30(5): 19-24.

Prochaska JJ, Sallis JF, Slymen DJ, et al. (2003) A longitudinal study of children's enjoyment of physical education. *Pediatric Exercise Science* 15(2): 170-178.

Prusak KA and Vincent SD (2005) Is your class about something? Guiding principles for physical education teachers. *Journal of Physical Education, Recreation & Dance* 76(6): 25-28.

Ramírez-Vélez R, Rodrigues-Bezerra D, Correa-Bautista JE, et al. (2015) Reliability of Health-Related Physical Fitness Tests among Colombian Children and Adolescents: The FUPRECOL Study. *PLoS One* 10(10): 1-12.

Ruiz JR, Castro-Piñero J, Artero EG, et al. (2009) Predictive validity of health-related fitness in youth: A systematic review. *British Journal of Sports Medicine* 43(12): 909-923.

Ruiz JR, Castro-Piñero J, España-Romero V, et al. (2011) Field-based fitness assessment in young people: The ALPHA health-related fitness test battery for children and adolescents. *British Journal of Sports Medicine* 45(6): 518-524.

Silverman S (2017) Attitude research in physical education: A review. *Journal of Teaching in Physical Education* 36(3): 303-312.

Simonton KL, Mercier K and Garn AC (2019) Do fitness test performances predict students' attitudes and emotions toward physical education? *Physical Education and Sport Pedagogy* 24(6): 549-564.

Solmon MA (2003) Student issues in physical education classes: Attitudes, cognition, and motivation. *Student learning in physical education: Applying research to enhance instruction* 2: 147-164.

Taylor-Powell E and Renner M (2003) Analyzing qualitative data. University of Wisconsin-Extension, Cooperative Extension.

Thompson HR, Linchey JK, King B, et al. (2019) Accuracy of School Staff-Measured Height and Weight Used for Body Mass Index Screening and Reporting. *Journal of School Health* 89(8): 629-635.

UNESCO (2014) *World-wide survey of school physical education.* Paris: UNESCO Publishing.

Vazou S, Mischo A, Ladwig MA, et al. (2019) Psychologically informed physical fitness practice in schools: A field experiment. *Psychology of Sport and Exercise* 40: 143-151.

WHO (2018) *Global action plan on physical activity 2018–2030: more active people for a healthier world.* Geneva: World Health Organisation.

578 Wiersma LD and Sherman CP (2008) The Responsible Use of Youth Fitness Testing to
579 Enhance Student Motivation, Enjoyment, and Performance. *Measurement in*
580 *Physical Education and Exercise Science* 12(3): 167-183.
581 Wrench A and Garrett R (2008) Pleasure and pain: Experiences of fitness testing.
582 *European Physical Education Review* 14(3): 325-346.
583 Zhu X, Chen S and Parrott J (2014) Adolescents' interest and performances in aerobic
584 fitness testing. *Journal of Teaching in Physical Education* 33(1): 53-67.
585 Zhu X, Davis S, Kirk TN, et al. (2018) Inappropriate Practices in Fitness Testing and
586 Reporting: Alternative Strategies. *Journal of Physical Education, Recreation &*
587 *Dance* 89(3): 46-51.
588
589

590