The Physical Educator Vol. 76 • pp. 634–660 • 2019
https://doi.org/10.18666/TPE-2019-V76-I3-8850

ASSESSMENT

Early Validation Evidence of the Canadian Practitioner-Based Assessment of Physical Literacy in Secondary Physical Education

Ken R. Lodewyk

Abstract

The assessment of physical literacy is vital and challenging for researchers and practitioners. Passport for Life is a Canadian global practitioner-based, formative, criterion-referenced assessment of physical literacy that is recorded in an online program. Students complete eight assessments divided across four broad categories of physical literacy (active participation, fitness, movement, and living skills). This study is an endeavor to uncover initial validation and feasibility evidence for the proposed use of Passport for Life with Grade 10–12 physical education students from various regions of Canada. The Standards for Educational and Psychological Testing were used as the theoretical framework. The results with 1,003 secondary school students provide score validity evidence on the content, format, and administrative guidelines of the items and scales; response process evidence about how the program was valued and how the assessments were implemented; internal structure evidence about how participants performed and the extent of associations between items within each scale; predictive evidence as to how scales related temporally to a second measurement of each assessment scale; and concurrent evidence about how the scales related to those of other components. While drawing attention to certain limitations and adding some recommendations,

Ken R. Lodewyk is an associate professor, Department of Kinesiology, Brock University. Please send author correspondence to klodewyk@brocku.ca

this article shows that the results generally support the Grade 10–12 Passport for Life relative to its intended use.

Physical literacy is "the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for engagement in physical activities for life" (International Physical Literacy Association, 2015). This concept has received increasing conceptual, professional, empirical, and curricular attention and development, particularly in the domains of sport (Jurbala, 2015; Lundvall, 2015) and physical education (PE; Edwards, Bryant, Keegan, Morgan, & Jones, 2017; United Nations Educational, Scientific, and Cultural Organization, 2015). Most of these developments stem from the conceptual work of Whitehead (2010), who envisioned physical literacy to be a holistic embodied experiential personal journey toward active, healthy, and meaningful being. Similar conceptualizations have been adopted by a variety of national PE associations including Physical and Health Education Canada (PHE Canada, 2015) and the Society of Health and Physical Educators (SHAPE America, 2014). Perhaps most for reasons of feasibility and policy, many of these initiatives have lessened emphasis on Whitehead's (2010) monistic, contextualized, and embodied movement experiences (Lundvall, 2015) to highlight elements of a fit and healthy lifestyle (Longmuir & Tremblay, 2016), a competent movement foundation (Silverman & Mercier, 2015), relevant and useful cognitive knowledge (Ennis, 2015), and availing affect (Chen, 2015; Weiss, 2011), which are primary aims of PE curricula across Canada (Kilborn, Lorusso, & Francis, 2015) and elsewhere (Dudley, 2015; SHAPE America, 2014).

The assessment of physical literacy has also received considerable attention as it is vital and poses particular challenges to researchers and practitioners because of the enormity of its scope and its varied conceptualizations, multidimensionality, fluidity, and proportion of irreducible qualities (Corbin, 2016; Giblin, Collins, & Button, 2014). Despite these challenges, the assessment of physical literacy is vital, should be authentic and feasible for practitioners, and needs to closely align with one's context and applied definition of physical literacy (Corbin, 2016). Several Canadian attempts include the Canadian Assessment of Physical Literacy (CAPL), Physical Literacy Assessment for Youth (PLAY Tools), and the Passport for Life

(PFL). In their review of these, Robinson and Randall (2017) note that despite compelling evidence for PFL's usability, applicability to practice, and fidelity to Whitehead's (2010) physical literacy ideals, more peer-reviewed evidence for its trustworthiness is yet needed. There is some validation evidence, that is, "the degree to which evidence and theory support the interpretation of test scores," for its proposed use (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014)—with Grade 3–9 PE students in certain regions of Canada (Lodewyk & Mandigo, 2017). The aim of this study was to uncover similar initial evidence for the use of the PFL with Grade 10–12 PE students in Canada and, if applicable, to add to the considerable support for its feasibility (usability).

PFL was developed specifically by PE experts to authentically promote, develop, and assess physical literacy in K–12 PE students across Canada while serving as a useful online resource for PE teachers (see http://passportforlife.ca/). To fulfill this aim, they designed PFL to emphasize feasibility (i.e., clarity, attainable, online recording, economical, efficient, safe, and timely) over elite psychometrics, timely duration to maximize participation and retention, and criterion (to serve as subgoals on one's journey) over norm-referencing (comparison to a population), and to be practitioner friendly instead of relying on highly trained technicians to administer the assessments. Other vital PFL aims include ensuring the assessments are "inclusive, adaptable, individualized, relevant (e.g., aligned to curriculum), process (versus product) oriented, safe, comprehensive, ongoing, affirming, valid, efficient and clear, useful for formative (learning) rather than summative (grades) and for the promotion of life-long development of physical literacy" (Lodewyk & Mandigo, 2017, p. 445). In endeavoring to meet these aims while reflecting a holistic conceptualization of students' physical literacy, the self-report assessments of PFL often had a single or a few items for some scales (e.g., cooperation, moderate-to-vigorous physical activity [MVPA], self-efficacy, empathy, perspective-taking). The advantages of this are clear and likely outweigh somewhat diminished psychometric properties that may result from using shortened scale assessments (Gosling, Rentfrow, & Swann, 2003; Horvath & Röthlin, 2018).

Extensive consultation and planning led to the selection of four PFL components of physical literacy known as active participation, movement, fitness, and living skills. The initial version was developed for and implemented with Grade 3–9 PE classes (for a review, see Lodewyk & Mandigo, 2017). Shortly thereafter (2014–2015), the Grade 10–12 program was piloted and then full implementation began in 2015–2016. In each version, students first complete a short demographic survey and then complete eight assessments: one for active participation (self-report survey of participation, interests, and preferred environments for physical activities), three fitness skills (core strength, cardiorespiratory endurance, and dynamic stability), three movement skills (locomotion, object control, and object manipulation), and one for living skills (a self-report survey assessing feeling, thinking, and relating). The teacher administers each assessment during PE—with the option of having students complete the online surveys for homework—and can enter and manage resources and student data in a secured online platform housed and maintained by PHE Canada. This enables teachers to track students' results over time so they can provide more specific attention to students and classes where necessary. Their online teacher resource also provides them information about how to interpret students' PFL results and help them set goals and work for ongoing improvement. Rather than being given a score for each component, students receive a personal paper "passport" that summarizes their results on each completed assessment and provides suggestions for how they might further develop on the components.

This study was theoretically grounded in the broad guidelines of the Standards for Educational and Psychological Testing (Standards; AERA, APA, & NCME, 2014). While traditional validation processes have focused on demonstrating a variety of types of validity (e.g., content, criterion, and construct) to validate an instrument, the Standards recommend the provision of ongoing evidence about the relative validity of the applications of an instrument based on properties of the test scores and contextual interpretations derived from particular uses of the instrument. The Standards further recommend the provision of multiple forms of validity evidence (content, response processes, internal structure, relations with other variables, and consequences of testing) to arrive at a more global perspective

of the validity evidence. This study provides score validity evidence based on (1) the content, format, and administrative guidelines of its items and scales; (2) response process evidence about how the assessments were understood, implemented, and performed; (3) internal structure evidence about how the items relate to their scale; (4) predictive evidence as to how scales relate temporally to a second measurement of each assessment scale; and (5) concurrent evidence about how the scales relate to scales of other components and how the four components (i.e., active participation, fitness, movement, and living skills) are associated, including to daily physical activity and MVPA.

Method

Participants

The sample size was 1,003 ($n_{females}$ = 431 or 43% ; n_{males} = 572 or 57%) in Grades 10 (n = 496, 49.5%), 11 (n = 227, 22.6%), and 12 (n = 263, 26.2%); 17 students (1.7%) did not report grade. Most (90%) were from the provinces of Manitoba (n = 519, 51.7%) and Ontario (n = 380, 37.9%) with the remainder from Alberta (n = 35), British Columbia (n = 16), Saskatchewan (n = 11), and Nova Scotia (n = 3). Several (n = 39) did not report their province. The quantities of participants completing each dimension by Assessment Time 1 and 2, respectively, were 818 and 122 for active participation, 707 and 88 for living skills, 523 and 148 for fitness skills, and 558 and 127 for movement skills. The living skills sample was reduced to 707 from 730 following a screening of multivariate outliers based on protocol (i.e., excessive Mahalanobis distance values; Tabachnick & Fidell, 2006).

Data Collection

As in prior uses of PFL (Grades 3–9), students' assessment data were entered into the password-protected online PFL database hosted by PHE Canada, by participating teachers who received a password and username to enter the program, register their class, and enter data for any of the assessments. For each class and/or student, teachers could then request Passports that provided concise yet useful formative feedback on students' physical literacy. Since there were no individual identifiers (names are replaced with a

unique code for each student), students were given a password and selected a username to access PFL so they could complete the online self-report surveys and view the summary and explanation of the assessments and their results on them. A university research ethics board and by PHE Canada approved the secondary use of the data (Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada, 2014) for the purposes of this study.

Measures

Feedback from practicing teachers. Of the 26 participating teachers in the pilot study of PFL Grade 10–12, 14 voluntarily completed an online Exit Survey (12 did not due to their work-to-rule status at the time) that was developed by PHE Canada program managers to provide useful teacher feedback about the assessments, the online system, the teacher resources, and whether the PFL helps teachers to facilitate physical literacy development in their students. The survey consisted of 55 items responded through both a 4-point Likert scale (1 = *strongly agree*, 4 = *strongly disagree*) and some written responses.

Student profile. Each participant was asked to provide some demographic information (e.g., gender, grade level) and responses to two questions when they initially entered the online PFL website. The two questions were "How many times a week do you have PE class?" with response choices from 0 to 5, and "How long are you physically active each day?" with options being less than 20 min (1), 20 to 40 min (2), 41 to 60 min (3), or more than 60 min (4).

Active participation. The active participation component consisted of 46 self-report items within these nine categories: MVPA (3 items), perceived fitness level (1 item), hours of daily sleep time (2 items), hours of daily sedentary screen time (2 items), barriers to physical activity (9 items), participation in diverse physical activities (19 items) and environments (5 items), and interest in participating in more diverse physical activities (4 items). The daily physical activity item from the Student Profile Survey was also included in the active participation component. Two other active participation items were not within the purposes of this study. Category scores consisted of the mean of all of the items within that category. Each

item was designed to assess a relatively unique aspect of active participation, so items within the categories were not expected to be positively and significantly correlated within categories and alpha reliability was not relevant or computed for the active participation categories.

Many (n = 18) of the active participation items were also used in the Grade 7–9 PFL survey of active participation. Of these, eight items assessed participation in the diverse activities of team sports, individual sports and physical activities, dance, and fitness activities at home and at school (a list of sample activities was included); four items assessed participants' interest in participating in more diverse physical activities; one item assessed daily physical activity; and five items measured level of participation in physical activities on or in land, snow or ice, water, air, and outdoor settings. Response choices were on a 4-point scale generally ranging from *strongly disagree* (1) to *strongly agree* (5).

The remaining items (n = 28) in the PFL Grade 10–12 active participation survey consisted of two items each for sleep, sedentary screen time, and MVPA that were adapted from previous uses (Hay & Cairney, 2006). For example, the sedentary screen time item for schools days was "On most typical school days (Monday to Friday), I spend approximately the following number of hours per day inactive and in front of a screen (e.g., phone, TV, video, computer, video game) during my own free time." It offered students 12 response choices (0–Over 10). This was also the response choices for the two MVPA items with a similar question stem, whereas the sleep items responses could range from 0–Over 12. The third MVPA item ("How many days in the last week did you exercise/participate in moderate-to-intense physical activity for at least 20 minutes?") provided eight response choices (0–Everyday) and was adapted from the Active Australia Survey (Australian Institute of Health and Welfare, 2003). Finally, the perceived fitness level item has been used previously (e.g., Lodewyk & Sullivan, 2015), used a 5-point response scale (1 = *very poor*, 5 = *very good*), and asked students, "Which of following most closely describes your current level of physical fitness (strength, endurance, flexibility…)?"

The author(s) of the survey (n = 20) developed items for nine research-based barriers to physical activity (work, technology,

weather, season, peers and lack of time, finances, parental/guardian support, and transportation) that used the stem "Being regularly physically active is difficult for me because" and a 5-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*). The designers also included 11 items assessing activities of daily living that were part of the participation in diverse activities subscale. These were walking or bicycling to and/or from school, walking or bicycling to and/or from work, walking or bicycling elsewhere (e.g., post office), household chores (vacuuming, cleaning, doing dishes…), yard work (raking leaves, lawn mowing, shoveling snow), gardening, washing clothes, active hobbies (crafts, sewing, woodwork, auto mechanics), baking/cooking, shopping (for home and personal essentials), and being physically active as an employee at work.

Living skills. The living skills component consisted of 49 items related to physical literacy that assessed feelings (17 items), thinking (20 items), and relating (12 items). Items were scored on a 4-point scale (1 = *never*, 2 = *sometimes*, 3 = *most of the time*, 4 = *all of the time*). Students with higher positive feeling tend to report that physical activity and learning new physical activities is important, interesting, and useful to them. They have a high level of confidence to successfully perform physical activity and physical education and are generally satisfied with their physical appearance. They also report higher positive emotional experiences during physical activity (e.g., alert, excited, determined, enthusiastic, enjoyable, and proud) and lower negative feelings such as being irritable, distressed, upset, nervous, afraid, and embarrassed in physical activity settings. Those reporting high scores on the thinking survey are more likely to intend to exercise regularly and do so because of more intrinsic reasons such as a life goal or choosing to be healthier; know about and how to monitor activity intensity, cultural diversity, the benefits of regular physical activity, eating and sleeping well, and managing stress; report that one's physical ability can be improved with effort, learning, and practice; carefully consider and apply previous insight, evidence, and different possibilities to make a difficult decision; are able to regulate obstacles to their attention, learning, and motivation to persist in physical activity; are able to set challenging and realistic short- and long-term standards to aim to achieve; and report managing their activity, time, peers, and sources of help to

enhance their being physically active. Finally, students with higher averages for interacting tend to be appropriately assertive, listen actively, apply self-control, have friends, and cooperate (get along with, compliment, and respect others). They are also more empathic, as they report having legitimate feelings of compassion and concern for someone who is encountering some difficulty or misfortune. They also try to understand another person's perspective or feelings compared to their own before coming to conclusions or making decisions.

Of the 49 items, 20 were from the existing Grade 7–9 PFL living skills survey with minor modifications for the Grade 10–12 level. The 29 additional items were used previously in other validated measures and modified for relevance to physical literacy (i.e., regular physical activity and health). They consisted of an additional item each for self-efficacy, value, critical thinking, peer learning, effort regulation, and goal setting from the Motivated Strategies for Learning Questionnaire (Pintrich, Smith, Garcia, & McKeachie, 1991); positive (6) and negative (6) feelings from the Positive and Negative Affective States survey (Watson, Clark, & Tellegen, 1988), with the latter being recoded to a positive valence for this study; two perspective-taking items and one empathy item from the Interpersonal Reactivity Index (Davis, 1983); an incremental ability conception item from the Conceptions of the Nature of Athletic Ability II Questionnaire (Wang, Liu, Biddle, & Spray, 2005); two autonomous motivation items from the Treatment Self-Regulation Questionnaire (Williams, Freedman, & Deci, 1996); a cooperation item from the Modified Aggression Scale (Orpinas, 1993); and two knowledge items from the Active Australia Survey (Australian Institute of Health and Welfare, 2003). The two remaining items were newly developed by the survey authors and assessed active listening within the related subscale ("I am able to easily listen to someone who is talking to me about something important without interrupting them") and understanding within the thinking subscale ("I am aware of how other cultures view the importance of the body and movement compared to mine").

Fitness skills. The teacher guides, tasks, and scoring rubrics and scale (1 = *emerging*, 2 = *developing*, 3 = *acquired*, 4 = *accomplished*) for the three fitness skill assessments closely resembled those for

Grades 7–9. The fitness assessment that differed most was the agility and balance assessment that involved a "hexagon hop" rather than the previously used "lateral bound." The aim of this task was to assess each student's ability to consistently maintain a state of balance while changing direction with speed. The comparison of clockwise and counterclockwise trials could also reveal imbalance between the right and left movement skills. The task basically involved making five small hexagons on the floor using tape with two students (one participant and one observer) at each. The teacher determines and informs the participant of the direction of the hops (clockwise or counterclockwise) and the foot to be used for hopping for each of four trials and prompts them when to begin. The participant begins in the center of the hexagon facing forward and will rapidly complete sequences of hops out of and back into the hexagon, landing on both feet and maintaining a forward-facing position without stopping throughout the sequence (until returning to the starting position). Each student performs the assessment in both directions—clockwise and counterclockwise—for a total of 2 sequences with a 1-min rest period between each.

The aim and general protocol of the second fitness skill (core strength) for the "plank challenge" remained as they were in Grades 7–9. The plank challenge was altered so students could choose a more difficult plank position compared to the regular and least difficult four-point ("table-top") position used in Grades 7–9. If they chose a higher (medium) level of difficulty, they performed a three-point plank, which was the same as the initial (least difficult) plank except students then extended one arm or foot parallel to the floor and fully extended. The most difficult level option involved a two-point plank wherein they initially set up in the regular plank position, then extend one foot and the opposite arm parallel to the floor and fully extended. To account for these differences in difficulty level, the second (medium, three-point plank) level served as the average so the data for students in that level was not altered. One point (to a minimum of 1) was subtracted from those who chose to complete the first (easiest) level, whereas 1 point was added (to a maximum of 4) to those selecting the third (most difficult) level. The 60-s trial (followed by a second trial) begins when the correct plank position is performed.

The aim of the third fitness skill (cardiorespiratory endurance) for the "four-station circuit" remained the same as for Grades 7–9. The students performed a different movement pattern continuously for 30 s at each station prior to rotating to the next station. The overall duration was changed for Grades 10–12 to enable participants the option of continuing beyond 12 min to a maximum of 20 min. In addition, each circuit station and scoring rubric was revamped. The first station remained the "agility ladder" except the foot pattern changed from a two-foot hop in Grades 7–9 to students hopping on one foot in the following pattern: dominant foot in, nondominant foot in, dominant foot out, nondominant foot out over the next square of ladder. They continued the "one in, one in, one out, one out" pattern up the ladder and then turned to go back (returning) through the ladder. The second station consisted of "mountain climbers" wherein students start in the hand-plank position and, while maintaining good core stability, flex one knee toward the chest, then rapidly extend that knee and simultaneously flex the opposite knee toward the chest. "Cone carousel" was the name of the third station. Students begin at the first of the three pylons arranged in a triangular pattern, while facing forward in the direction of travel. They move forward quickly to the second pylon, shuffle quickly to the third pylon, then backpedal quickly to the first pylon. The final station was called "split jumps/push" wherein students move from a squat to a standing position with feet parallel and shoulder width apart and arms elevated overhead. On the teacher's signal they (in one fluid motion jump while maintaining hands elevated overhead) split their feet, placing one foot on either side of the badminton end line, and immediately return to the start position with feet parallel. Without pause, the students jump split again, placing their other foot forward while maintaining the hands elevated overhead.

Movement skills. The teacher guides, tasks, and scoring rubrics and scale (1 = *emerging*, 2 = *developing*, 3 = *acquired*, 4 = *accomplished*) for the three movement skill assessments also closely resembled those for Grades 7–9 except for modifications to increase the difficulty. Each student completes two trials of each assessment. The locomotion assessment was changed from "run, side shuffle, backpedal" in Grades 7–9 to "run, cross over, shuffle, backpedal" in Grades 10–12. The setup involves setting up two pylons 10 m apart

from each other with the first pylon (A) serving as the starting point and the other pylon being B. Two other pylons (C and D) are set up 5 m from and at 45 degrees from Pylon A in the direction of and each flanking Pylon B. All pylons should be a minimum of 3 m from any outside wall. On the tester's prompt, the student runs forward as fast as he or she can and slightly past Pylon B (10 m). At Pylon B, the student then performs an appropriate crossover step (right leg crosses over left to travel to the left) on a 45-degree angle to Pylon C (5 m). The student cuts around Pylon C and then performs a rightward side shuffle to return to the outside of Pylon B (5 m). The student then performs an appropriate crossover step (left leg crosses over right to travel to the right) on a 45-degree angle to Pylon D (5 m). The student cuts around Pylon D and then performs a leftward side shuffle to return to Pylon B (5 m). The student goes around Pylon B and backpedals toward Pylon A (10 m), completing the task. Using the teacher rubric for guidance, the instructor observes for strong acceleration and deceleration, control, and fluid conversion between running, crossing over, shuffling, and backpedaling.

Dynamic running movement was added to the start of both the object control ("throw and catch") and object manipulation ("punt and catch") movement skill assessments used in Grades 7–9, hence the new titles for each ("run, throw, and catch" and "run, punt, and catch"). The floor setup for the "run, throw, and catch" was similar except for having three lines parallel to the wall at 2.0 m, 2.75 m, and 3.0 m from the wall with the 1 m wide target box running between the 2.0-m and 2.75-m lines. A taped spot on the floor 4 m behind the 3-m throwing line serves as a starting point for the seated student. On the prompt to start, the student stands, turns, and travels 4 m (to behind the 3-m line) and throws the ball at the wall so it bounces in the target area *after* hitting the wall. The student then catches the ball after the bounce while remaining behind the 3-m line. The scoring rubric guides observers to additionally check (from Grades 7–9) whether the student is able to rise and turn without stumbling or falling, travels forward using a mature running pattern, moves in a continuous fluid motion toward the target, and crosses the throwing line (3.0 m) on the floor when throwing or catching. Finally, a short run was similarly added to the start of the object manipulation ("kicks" involving a punt and catch) Grade 7–9 assessment. In

the revised object manipulation ("run, punt, and catch") assessment, participants attempt to demonstrate a mature and controlled running pattern prior to coordinating hand and foot to control kick (punt) a ball and then to predict the motion of the ball so they successfully pursue, intercept, and catch it without crossing the restraining line 3 m from the wall. The scoring rubric was also adapted to add the relevant observations of fluidity, control, and maturity of running during the task.

Results

Content Evidence

Development. The development of the Grade 10–12 PFL involved an extensive consultation process that has been previously reported (Lodewyk & Mandigo, 2017). During and following the development and initiation of the Grade 3–9 PFL, a team of six domain experts that included two university professors met occasionally to plan, write, review, and edit iterations of the Grade 10–12 PFL. Once again, each member of this team represented a balance of expertise in each of PFL components from across Canada and had teaching-coaching experiences with adolescents in physical activity settings, and five had teaching certification in PE. The Grade 10–12 PFL assessments were also developed after a thorough review of the feedback about the Grades 3–9 PFL assessments from the data, advisory committee, teachers, parents, and students. In line with that feedback, several assessment items were altered and retained and many new items were added. Care was taken to ensure that the content and nature of the assessments was well-justified in the academic (research) literature relative to physical literacy and health outcomes.

Data from a pilot test of a draft of the Grade 10–12 PFL conducted a year prior to the finalized version in 2015–2016 was also considered as feedback in the development process. The pilot study was conducted with a sample of 642 students in Grades 10–12 (n_{females} = 261; n_{males} = 381) from four Canadian provinces (Alberta, Saskatchewan, Manitoba, and Ontario) the year before the data for this study. Relatively normal scale distributions, satisfactory-to-high alpha reliability (.66–.93), and significant relations in theoretically justified directions provided evidence for the interactive nature of

active participation, fitness, living, and movement skills for physical literacy, and for the general validity of the assessments for the intended use. For example, there were positive and significant correlations (.50–.84) between each living skill assessment (feeling, thinking, and interacting) and between each of the fitness and movement skill assessments (.23–.54), and between these skills and fitness level (.32–.45), MVPA (.27–.39), and participation and interest in fitness activities and team and individual sports (.18–.48). Through the feedback from this pilot test, minor modifications were made to the wording of some of the items.

Components. The four PFL assessment components with their scales (see Table 1) venture to reflect a holistic conceptualization of physical literacy (Dudley, 2015; PHE Canada, 2015; Roetert & MacDonald, 2015; Whitehead, 2010). For example, the Grade 10–12 PFL includes at least one assessment item for each concept to be included in a holistic understanding of physical literacy provided by Corbin (2016). Readers are also referred to Lodewyk and Mandigo (2017) for more about the similarity of items between the Grade 3–9 and 10–12 PFL assessments. The additional active participation items in the Grade 10–12 PFL reflect students' increased cognitive capability to complete longer surveys (Jordan & Turner, 2008), increased diversity and personalization of activities in high school PE curricula (Kilborn et al., 2015; Lacy & Hastad, 2003; SHAPE America, 2014), and the particular importance of psychosocial and environmental factors during adolescence when youth tend to enjoy more choices while likely facing added barriers to health and physical activity that require more self-regulatory skills (Chen, 2015; Weiss, 2011). For this reason, the Grade 10–12 active participation component added more assessment items for activities of daily living, sleep, sedentary screen time, MVPA, and barriers to physical activity. Additional items added to the living skills component included assessments of more specific feelings (e.g., afraid, determined), thinking skills (e.g., autonomous regulation, understanding, goal-setting), and relational skills (e.g., empathy, active listening, perspective-taking, and respect). Use of scales with only one or two Likert-style items might compromise some psychometric qualities while increasing feasibility for use in larger populations with limited time, but can also still be useful for use in empirical research (Horvath & Röthlin, 2018).

Table 1
Descriptive Statistics and Alpha Reliability for PFL Components by Assessment Time

	First assessment time			Second assessment time		
Scale (# of items)	***n***	***M* (*SD*)**	**α**	***n***	***M* (*SD*)**	**α**
Movement Skills (3)	558	2.90 (.77)	.75	127	3.50 (.58)	.69
1. Locomotion (1)	-	2.92 (.96)	-	-	3.67 (.58)	-
2. Object Control (1)	-	2.97 (.85)	-	-	3.43 (.77)	-
3. Object Manipulation (1)	-	2.80 (1.02)	-	-	3.39 (.85)	-
Fitness Skills (3)	523	2.67 (.79)	.64	148	3.33 (.67)	.66
1. Speed/Agility/Balance (1)	-	2.81 (.85)	-	-	3.35 (.80)	-
2. Core Strength (1)	-	2.58 (1.13)	-	-	3.31 (.87)	-
3. Cardiorespiratory Endurance (1)	-	2.63 (1.10)	-	-	3.34 (.94)	-
Living Skills (49)	708	3.10 (.39)	.94	88	3.17 (.44)	.94
1. Feeling (17)	-	3.08 (.49)	.89	-	3.18 (.43)	.85
2. Thinking (20)	-	3.14 (.42)	.91	-	3.15 (.42)	.90
3. Interacting (12)	-	3.06 (.43)	.83	-	3.17 (.44)	.84
Active Participation (46)	818	2.69 (.36)	-	122	2.84 (.35)	-
1. Daily Physical Activity (1)	853	3.33 (.84)	-	0	-	-
2. MVPA (3)	-	3.68 (1.78)	-	-	3.39 (1.06)	-
3. Fitness Level (1)	-	2.80 (.86)	-	-	3.09 (.75)	-
4. Sleep (2)	-	8.17 (1.30)	-	-	8.52 (1.41)	-
5. Sedentary Screen Time (2)	-	4.57 (2.30)	-	-	4.20 (2.17)	-
6. Barriers to Physical Activity (9)	-	2.29 (.36)	-	-	2.31 (.38)	-
7. Participation in Diverse PA (19)	-	2.24 (.44)	-	-	2.38 (.45)	-
8. Diverse Environments (5)	-	2.36 (.54)	-	-	2.50 (.56)	-
9. Interest in More Diverse PA (4)	-	2.47 (.61)	-	-	2.64 (.57)	-

Note. All on a 4-point scale except for daily physical activity, MVPA (moderate-to-vigorous physical activity), and sedentary screen time.

Response Process Evidence

Exit Survey data from participating teachers following their use of PFL in the piloted version that was almost identical to this version provided useful feedback about teachers' and their students' experiences with the assessments relative to the aim of assessing,

promoting, and developing physical literacy. All of the teachers were very positive about the way PFL helped to increase their students' understanding of physical literacy, noted that it was useful to their teaching, and said that items for active participation and living skills were easy for their students to follow and understand. The majority of teachers (85–95%) reported that the results of each assessment enabled their accurate evaluation of their students' physical literacy; that the teacher support resources were useful, relevant, and easy to apply; and that it was easy to access, register students, navigate, and enter and interpret the data and results in the online system (Web usability). A lower yet still high proportion of teachers (70–85%) reported that PFL would increase their knowledge of physical literacy, change aspects of their teaching, and help them to identify gaps in and improve their students' physical literacy. Teachers' lowest ratings (45%) were that PFL assessments—especially those for movement and fitness skills—took a reasonable amount of time to complete (on average 4–6 classes). Several teachers also recommended adding more video aids, which PHE Canada has since done (T. Zakaria, personal communication, December 16, 2017).

Internal Structure Evidence

The descriptive (i.e., means and standard deviations) and normality (i.e., skew and kurtosis) statistics did not signal any discrepancies for a sample of this size (Tabachnick & Fidell, 2006). The Standards (AERA, APA, & NCME, 2014) also note the importance of (relative to an assessments' intended interpretation) assessment item scores aligning with the construct being assessed. The factor structure of each living skills dimension (feeling, thinking, and relating) was explored via a separate principal component exploratory factor analysis with only the first assessment data since the sample size of the second assessment period was insufficient (< 200; Tabachnick & Fidell, 2006). In line with theoretical expectations for each dimension, each analysis requested a single factor and suppressed factor loadings < .30. Results revealed that except for one thinking (knowledge) and one feeling (negative affect) item, each item had a satisfactory-to-high factor loading (.30–.81), the extracted factor explained a satisfactory proportion of the variance (feeling = 38.49%, thinking = 36.76%, relating = 35.78%), and the eigenvalues were 6.54 for feeling, 7.35 for thinking, and 4.29 for relating. The two items

that failed to adequately load onto their factor were omitted from subsequent analyses in the study. The alpha reliability coefficients for each of the extracted living skills dimensions were high for both assessment times (α = .83–.91), as was the alpha reliability for the 47 living skills items (α = .94).

Pearson bivariate correlations were computed between each scale within each component (movement, fitness, living skills, and active participation) for the first assessment period. The results (Tables 2 and 3) reinforced expected scale associations within each component from previous research based on the level and direction of the correlations and the absence of multicollinearity. There were positive and moderate-to-high correlations between the three movement skills (r = .46–.55), fitness skills (r = .36–.43), living skills (r = .46–.78), and most of the active participation scales as only five of the 36 active participation relationships were not significant (p = .01). For example, as sedentary screen time and perceived barriers to physical activity decrease, most of the remaining aspects of active participation tend to increase (daily and diverse physical activity, MVPA, perceived fitness level, diverse environments, and interest in diverse physical activities). These results, the satisfactory evidence of the factor structure of the three living skills dimensions (feeling, thinking, and relating), and sound alpha reliability of relevant scales support the internal structure of each PFL component and scale.

Table 2
Component and Physical Activity Correlations

Scale (# of items)	1	2	3	4	5
1. Movement Skills (3)	-				
2. Fitness Skills (3)	.60*	-			
3. Living Skills (49)	.32*	.29*	-		
4. Active Participation (46)	.37*	.36*	.57*	-	
5. Daily Physical Activity (1)	.34*	.33*	.36*	.44*	-
6. MVPA (3)	.27*	.29*	.42*	.62*	.47*

Note. MVPA = moderate-to-vigorous physical activity.

*p < .01. Time 1 only.

Table 3

Component Scale Correlations

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. MS1 (Locom.)	-																
2. MS2 (Obj. Con.)	.46**	-															
3. MS3 (Obj. Man.)	.52**	.55**	-														
4. FS1 (SAB)	.47**	.42**	.41**	-													
5. FS2 (Core Str.)	.45**	.27**	.28**	.43**	-												
6. FS3 (Card. End.)	.43**	.36**	.35**	.36**	.37**	-											
7. LS1 (Feeling)	..32**	.28**	.32**	.37**	.22**	.24**	-										
8. LS2 (Thinking)	.20**	.19**	.21**	.22**	.16**	.19**	.67**	-									
9. LS3 (Interacting)	.13*	.13*	.10	.11*	.09	.07	.46**	.78**	-								
10. AP1 (DPA)	.32**	.27**	.24**	.24**	.27**	.25**	.38**	.31**	.21**	-							
11. AP2 (MVPA)	.41**	.24**	.25**	.27**	.23**	.30**	.39**	.41**	.34**	.55**	-						
12. AP3 (Fitness)	.45**	.29**	.32**	.33**	.40**	.44**	.53**	.43**	.26**	.45**	.59**	-					
13. AP4 (Sleep)	.09	.07	.05	.11*	.05	.06	.11	.09*	.05	.08*	.15**	.15**	-				
14. AP5 (Screen)	-.20**	-.16**	-.10*	-.12*	-.14**	-.20**	-.28**	-.22**	-.16**	-.19**	-.27**	-.23**	-.03	-			
15. AP6 (Barriers)	-.19**	-.24**	-.20**	-.22**	-.16**	-.21**	-.29	-.14**	-.05	-.27**	-.31**	-.29**	-.06	.25**	-		
16. AP7 (Div. PA)	.19**	.16**	.17**	.11*	.17**	.10*	.29**	.44**	.39**	.27**	.43**	.36**	.14**	-.15**	-.08*	-	
17. AP8 (Div. Env.)	.31**	.25**	.17**	.25**	.24**	.20**	.38**	.41**	.34**	.29**	.52**	.40**	.15**	-.29**	-.22**	.60**	-
18. AP9 (Interest)	.21**	.13**	.15**	.20**	.26**	.16**	.45**	.42**	.32**	.25**	.35**	.36**	.10*	-.22**	-.11**	.49**	.41**

Note. MS = movement skill; FS = fitness skill; LS = living skill; AP = active participation; Locom. = locomotion; Obj. Con. = object control; Obj. Man. = object manipulation; SAB = Speed/Agility/Balance; Str. = strength; Card. End. = cardiorespiratory endurance; DPA = daily physical activity; MVPA = moderate-to-vigorous physical activity; Div. PA = participation in diverse physical activities; Div. Env. = participation in diverse environments; Interest = interest in more physical activities.

*$p < .05$. **$p < .01$. Time 1 only.

Temporal Consistency

Stability of measurement across the two repeated assessment times (fall and spring seasons) was assessed with alpha reliability coefficients > .50 as the criterion based on McCrae's (2015) justification, particularly for scales with very few items. This level was also set since scales are likely to differ somewhat by assessment period because of the different seasons (fall and winter) in which they were administered. The results revealed that each scale met the criterion as the scales for active participation and living skills had coefficients ranging from .68 (the thinking dimension of living skills) to .88 (interest in more diverse activities), whereas those for movement and fitness skills were between .53 (object control) and .74 (core strength/plank). The coefficients between assessment times for the mean of the three movement skills ($\alpha = .61$), three fitness skills ($\alpha = .80$), three living skills ($\alpha = .73$), and active participation items ($\alpha = .89$) also signaled satisfactory stability across assessment periods.

Relations to Other Variables

The Standards (AERA, APA, & NCME, 2014) also report that another important form of validation evidence is the degree that scales of an assessment relate concurrently between components. As is evident in Table 3, correlations between the movement, fitness, and living skills were positive and significant ($p = .01$) except for those involving the relating scale of living skills. Further, except for the sleep scale, 67 of the 72 bivariate correlations between the active participation scales and either fitness, movement, or living skills were significant ($p = .01$) and in the theoretically expected direction. These relationships were also evident in the positive, moderate ($r = .29–.60$), and significant ($p = .01$) bivariate correlations between the composite means for active participation and movement, fitness, and living skills (Table 2). More specifically, except for some relationships—particularly those involving relating and sleep—as one scale or component of PFL increases or decreases, the more likely the others will also do so. Finally, correlation and regression analyses tested whether the different PFL components predicted two important physical literacy outcomes (levels of MVPA and daily physical activity). The results revealed moderate, positive, and significant ($p = .01$) bivariate correlations between both of these outcomes and

each scale (except for sleep) and component of PFL (Tables 2 and 3). MVPA was omitted from the active participation mean for the regression analyses. Results revealed that the four PFL components collectively predicted self-reported daily physical activity, $R^2 = .20$, $F(4, 351) = 21.94$, $p < .001$, and MVPA, $R^2 = .28$, $F(4, 359) = 34.66$, $p < .001$. Except for the prediction of MVPA by the overall mean of movement skills ($p = .76$), the PFL components each predicted MVPA (fitness skills, $p = .01$; living skills, $p = .01$; active participation, $p = .01$), and daily physical activity (movement skills, $p = .02$; fitness skills, $p = .01$; living skills, $p = .04$; active participation, $p = .001$).

Discussion

PFL is intended to serve as a feasible, formative, criterion-referenced, and practitioner-based assessment of a holistic physical literacy. The aim of this study was to discover any validation evidence for early use of the PFL with Grade 10–12 PE students in Canada. Theoretically grounded in the broad guidelines of the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 2014), the study provides score validity evidence on the content, format, and administrative guidelines of the items and scales; response process evidence about how the program was valued and how the assessments were implemented; internal structure evidence about how participants performed and the extent of associations between items within each scale; predictive evidence as to how scales related temporally to a second measurement of each assessment scale; and concurrent evidence about how the scales related to those of other components. While drawing attention to certain limitations and adding some recommendations, this article shows that the results generally support the Grade 10–12 PFL for its intended use.

The study revealed content validity evidence in the development of the Grade 10–12 PFL that was based on a thorough review of a pilot study and previous use of PFL for Grades 3–9, extensive consultations and planning by a balanced team of experts, and research-based justification of assessment items attempting to reflect a relatively holistic conceptualization of physical literacy (Dudley, 2015; Jurbala, 2015; Lloyd, 2016; Whitehead, 2010) that is aligned with many PE curricular standards and outcomes (Ontario Ministry of Education, 2015; PHE Canada, 2015; Roetert & MacDonald, 2015). The level and direction of the correlations between assessment items within

their scale or construct signaled alignment with the intended interpretation of each construct or scale. For example, there were positive and moderate-to-high correlations between the three assessment items within each of the movement, fitness, and living skill components and almost all of the active participation scales. Satisfactory factor structure of the three living skills dimensions (feeling, thinking, and relating) and sound alpha reliability of relevant scales also supported the internal structure of each PFL component and scale. Further, the assessment scales were relatively stable across the two repeated assessment times (fall and spring seasons) and also related as expected *between* components. Bivariate correlations between the assessments in one component (e.g., movement skills) and those in another component (e.g., fitness skills) were generally positive and significant. The different PFL components were also predictors of MVPA and daily physical activity.

Another important source of validation evidence was the feedback provided by participating teachers that generally signaled a positive experience with PFL, most notably that it facilitated understanding of physical literacy and that it was useful and easy enough to access, understand, administer, and interpret. Almost a third of the responding teachers had doubts that PFL would increase their knowledge of physical literacy, change aspects of their teaching, and help them to identify gaps in and improve their students' physical literacy. This reflects some research into PE teachers' views of physical literacy (e.g., Lynch & Soukop, 2016; Stoddart & Humbert, 2017) indicating confusion and misunderstanding of physical literacy and its role in PE and the need for an easing of teachers' concerns relative to understanding, applying, and assessing physical literacy. The results of this study signal that the Grade 10–12 PFL and its teacher resources might serve to facilitate this in physical educators.

A formidable challenge in the overall assessment of physical literacy is its duration. For example, in this study almost half of the responding teachers had concerns about the amount of class time required. This is likely to be an ongoing tension, particularly in the global assessment of physical literacy as a holistic (multidimensional), embodied, personal journey (Lundvall, 2015; Whitehead, 2010) because of its size and complexity. As PFL reflects, it is likely that some aspects of physical literacy are too contextual (e.g., curricular

variations for cultural or regional considerations) and abstractly embodied (e.g., aesthetic ability, communication, and creativity) to be assessed validly by practitioners in educational settings. Lundvall (2015) summarizes this concern:

> There are, however, skepticisms about the assessment of physical literacy. Skepticisms are around the legitimacy of whether the ideals expressed in the concept, such as empowerment, embodiment, values, identities, beliefs, and social relatedness, should be assessed as mechanically . . . In a school culture dominated by a performance code, it appears difficult to uphold the broad and inclusive definition of physical literacy which is characterized by an individual's potential and being in the world. Assessments related to measuring what separates people are no longer relevant and appropriate. Such measurements induce a risk of a maltreatment or misuse of the concept literacy in a goal-oriented and assessment-driven school system. (p. 116)

It is also important to acknowledge that any global assessment of physical literacy such as PFL will only partially reflect not precisely mirror it, and essentially involves "[identifying and defining] 'specific' characteristics of physical literacy and [developing] appropriate assessment procedures for each (as opposed to attempting to develop a 'general' test)" (Corbin, 2016, p. 21). The multiple components and scales of PFL reflect such an approach and provide evidence that a valid overall assessment of holistic physical literacy will require a significant duration of class time. Each of the four PFL components consisted of several essential assessments and the living skills and active participation surveys used subscales (e.g., physical activity, sedentary screen time, self-efficacy, cooperation) and often consisted of one or two Likert-style items. Although such reduction of scale items may compromise certain psychometric properties, it can and still be valid enough for use in empirical research and will increase the feasibility and authenticity for use with larger populations having limited time (Gosling et al., 2003; Horvath & Röthlin, 2018).

Since the Standards (AERA, APA, & NCME, 2014) emphasize the importance of a thorough understanding of the contextual purpose of any assessment, it is vital to recognize that PFL comprises a

set of authentic assessments, each designed to provide each student (and their teachers and parents) criterion-referenced information (not a precise summative quantity) about their progress relative to its four essential components of physical literacy to facilitate their reflection and long-term healthy active living. PFL includes the student, teacher, and parent in the information-sharing process, which is vital because many of the physical literacy outcomes are best met in a "holistic and cross-curricular way" that both the "school and its wider community can service" (Dudley, 2015, p. 238). The internal, concurrent, and predictive validity evidence from this study reinforces the sound evidence-based theoretical and structural foundation of PFL for its intended use with Grade 10–12 students in Canada (e.g., Chen, 2015; Corbin, 2016; Ennis, 2015; Giblin et al., 2014; Longmuir & Tremblay, 2016; Lundvall, 2015; Silverman & Mercier, 2015; Weiss, 2011).

Future research should continue to define and refine tools and tests for different aspects of physical literacy that can be added to an authentic large-scale assessment for learning programs such as PFL. As more longitudinal PFL data are gathered, increased analysis would be helpful into the concurrent and predictive relations, particularly by grade level or ethnic group, and to long-term healthy active living outcomes. Limitations of the study include potential confirmation bias, as the author was also involved in the design of several of the Grade 10–12 PFL assessments, and construct underrepresentation due to the low number of assessments and/or items representing each scale or component. For example, the Grade 10–12 PFL may need new or expanded assessments of some noted aspects of physical literacy such as movement functions, forms, and flows (Lloyd, 2016) and essential knowledge (Ennis, 2015), particularly in the form of rules, tactics, and strategies of movement (Dudley, 2015) since most learning outcomes in Canadian PE curricula are focused on movement skills, games, and sport technique (Kilborn et al., 2016). It would also be useful if the data sample provided more information on Grade 10–12 PFL participants with special needs and potential adaptations for them, along with more balanced representation of the constituent residents of Canada, particularly from northern regions. Nevertheless, this study provides important early validation evidence in favor of the use of PFL as an authentic,

inclusive, formative, practitioner-based, general assessment of relatively holistic physical literacy for Grades 10–12 PE in most regions of Canada.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association. https://doi.org/10.1177/001316446602600328

Australian Institute of Health and Welfare. (2003). *Indicators of health risk factors: The AIHW review.* Canberra, Australia: Author.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada. (2014). *Tri-council policy statement: Ethical conduct for research involving humans*. Ottawa, Canada: Government of Canada. https://doi.org/10.5962/bhl.title.59037

Chen, A. (2015). Operationalizing physical literacy for learners: Embodying the motivation to move. *Journal of Sport and Health Science, 4,* 125–131. https://doi.org/10.1016/j.jshs.2015.03.005

Corbin, C. (2016). Implications of physical literacy for research and practice: A commentary. *Research Quarterly for Exercise and Sport, 87*(1), 14–27. https://doi.org/10.1080/02701367.2016.1124722

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113–126. https://doi.org/10.1037//0022-3514.44.1.113

Dudley, D. A. (2015). A conceptual model of observed physical literacy. *Physical Educator, 72,* 236–260.

Edwards, L. C., Bryant, A. S., Keegan, R. J., Morgan, K., & Jones, A. M. (2017). Definitions, foundations, and associations of physical literacy: A systematic review. *Sports Medicine, 47,* 113–126. https://doi.org/10.1007/s40279-016-0560-7

Ennis, C. A. (2015). Knowledge, transfer, and innovation in physical literacy curricula. *Journal of Sport and Health Science, 4,* 119–124. https://doi.org/10.1016/j.jshs.2015.03.001

Giblin, S., Collins, D., & Button, C. (2014). Physical literacy: Importance, assessment, and future directions. *Sports Medicine, 44,* 1177–1184. https://doi.org/10.1007/s40279-014-0205-7

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37,* 504–528. https://doi.org/10.1016/s0092-6566(03)00046-1

Hay, J. A., & Cairney, J. (2006). Development of the Habitual Activity Estimation Scale for clinical research: A systematic approach. *Pediatric Exercise Science, 18,* 193–202. https://doi.org/10.1123/pes.18.2.193

Horvath, S., & Röthlin, P. (2018). How to improve athletes' return of investment: Shortening questionnaires in the applied sport psychology setting. *Journal of Applied Sport Psychology, 30,* 241–248. https://doi.org/10.1080/10413200.2017.1382020

International Physical Literacy Association. (2015). [Homepage]. Retrieved from https://www.physical-literacy.org.uk

Jordan, J. S., & Turner, B. A. (2008). The feasibility of single-item measures for organizational justice. *Measurement in Physical Education and Exercise Science, 12*, 237–257. https://doi.org/10.1080/10913670802349790

Jurbala, P. (2015). What is physical literacy, really? *Quest, 67*, 367–383. https://doi.org/10.1080/00336297.2015.1084341

Kilborn, M., Lorusso, J., & Francis, N. (2016). An analysis of Canadian physical education curricula. *European Physical Education Review, 22*, 23–46. https://doi.org/10.1177/1356336x15586909

Lacy, A. C., & Hastad, D. N. (2003). *Measurement and evaluation in physical education and exercise science* (4th ed.). Toronto, Canada: Pearson Education.

Lloyd, R. J. (2016). Becoming physical literate for life: Embracing the functions, forms, feelings, and flows of alternative and mainstream physical activity. *Journal of Teaching in Physical Education, 35,* 107–116. https://doi.org/10.1123/jtpe.2015-0068

Lodewyk, K. R., & Mandigo, J. (2017). Early validation evidence of a Canadian practitioner-based assessment of physical literacy in physical education: Passport for Life. *Physical Educator, 74*, 443–477. https://doi.org/10.18666/tpe-2017-v74-i3-7459

Lodewyk, K. R., & Sullivan, P. (2015). 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*, 601–615. https://doi.org/10.1080/17408989.2015.1095869

Longmuir, P. E., & Tremblay, M. S. (2016). Top 10 research questions related to physical literacy. *Research Quarterly for Exercise and Sport, 87*(1), 28–35. https://doi.org/10.1080/02701367.2016.1124671

Lundvall, S. (2015). Physical literacy in the field of physical education: A challenge and a possibility. *Journal of Sport and Health Science, 4,* 113–118. https://doi.org/10.1016/j.jshs.2015.02.001

Lynch, T., & Soukop, G. J. (2016). "Physical education", "health and physical education", "physical literacy" and "health literacy": Global nomenclature confusion. *Cogent Education, 3*, 1–22. https://doi.org/10.1080/2331186x.2016.1217820

McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review, 19*, 97–112. https://doi.org/10.1177/1088868314541857

Ontario Ministry of Education. (2015). *The Ontario curriculum Grades 1 to 8. Health and Physical Education*. Toronto, Canada: Queens Printer.

Orpinas, P. (1993). *Skills training and social influences for violence prevention in middle schools: A curriculum evaluation* (Unpublished doctoral dissertation). University of Texas Health Science Center at Houston, School of Public Health, Houston, TX.

Physical and Health Education Canada. (2015). What is physical literacy? Retrieved from http://www.phecanada.ca/programs/physical-literacy/what-physical-literacy

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)* (Technical Report No. 91-B-004). Ann Arbor, MI: University of Michigan, School of Education.

Robinson, D. B., & Randall, L. (2017). Marking physical literacy or missing the mark on physical literacy? A conceptual critique of Canada's physical literacy assessment instruments. *Measurement in Physical Education and Exercise Science, 21,* 40–55. https://doi.org/10.1080/1091367x.2016.1249793

Roetert, E. P., & MacDonald, L. C. (2015). Review: Unpacking the physical literacy concept for K–12 physical education: What should we expect the learner to master? *Journal of Sport and Health Science, 4*, 108–112. https://doi.org/10.1016/j.jshs.2015.03.002

Society of Health and Physical Educators. (2014). *National standards and grade-level outcomes for K–12 physical education*. Champaign, IL: Human Kinetics.

Silverman, S., & Mercier, K. (2015). Teaching for physical literacy: Implications to instructional design and PETE. *Journal of Sport and Health Science, 4,* 150–155. https://doi.org/10.1016/j.jshs.2015.03.003

Stoddart, A. L., & Humbert, M. L. (2017). Physical literacy is…? What teachers really know. *Phenex, 8*(3), 1–18.

Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics*. Boston, MA: Allyn & Bacon.

United Nations Educational, Scientific, and Cultural Organization. (2015). *Quality physical education*. Retrieved from http://unesdoc.unesco.org/images/0023/002311/231101E.pdf

Wang, C. K., Liu, W. C., Biddle, S. J., & Spray, C. M. (2005). Cross-cultural validation of the Conceptions of the Nature of Athletic Ability Questionnaire Version 2. *Personality and Individual Differences, 38*, 1245–1256. https://doi.org/10.1016/j.paid.2004.08.007

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

Weiss, M. R. (2011). Teach the children well: A holistic approach to developing psychosocial and behavioral competencies through physical education. *Quest, 63*, 55–65. https://doi.org/10.1080/00336297.2011.10483663

Whitehead, M. (2010). *Physical literacy throughout the life course*. London, United Kingdom: Routledge.

Williams, G. C., Freedman, Z., & Deci, E. L. (1996). Promoting motivation for diabetics' self-regulation of HbA1c. *Diabetes, 45*, (Suppl. 2), 13A.