Measuring physical literacy for an evidence-based approach: Validation of the French perceived physical literacy instrument for emerging adults

Joseph Gandrieau [a, b, *], Olivier Dieu [a], François Potdevin [a], Thibaut Derigny [a], Christophe Schnitzler [c]

[a] Univ. Lille, Univ. Artois, Univ. Littoral Côte d'Opale, ULR 7369 – URePSSS – Unité de Recherche Pluridisciplinaire Sport Santé Société, F-59000 Lille, France
[b] Université Côte d'Azur, LAMHESS, Nice, France
[c] Unité de Recherche Sport et sciences sociales, UR 1342, Université de Strasbourg, France

ARTICLE INFO

ABSTRACT

Background/Objectives: Physical Literacy (PL) is increasingly recognized as a key element in studies aimed at promoting Physical Activity (PA), but measurement tools for emerging adults and evidence for assessing and using this concept is lacking in a wide range of contexts. We aimed to validate the French version of the Perceived Physical Literacy Instrument (PPLI): a scale for investigating PL in young French adults.

Methods: After PL and PA data collection (n = 2,248, age_{mean} = 19 ± 1.53 yrs), exploratory and confirmatory factor analyses, Cronbach's α and Omega's ω and an Intraclass-Correlation analysis were undertaken. Spearman's rank correlation and the Boruta algorithm were used to investigate the association between PL and PA. Boruta's algorithm examined deeper external validation by analyzing the strength of an overall PL score in explaining PA, compared with separate dimensions of PL and individual characteristics (BMI, sex).

Results: Results showed an acceptable level of reliability (ICC = 0.91), internal validity (α = 0.88; ω = 0.77), and external validity (Rhô >0.18, p < 0.01). The Boruta algorithm highlighted that the construct of PL is a significant predictor of PA, although not the strongest one which is social and affective dimension.

Conclusion: This study provided data on validity and reliability of the first French assessment tool to measure PL constituted by four intertwined dimensions (physical, cognitive, social, affective). At the same time, it provides new evidence of the association between PL and PA.

1. Introduction

The concept of physical literacy (PL)[1] has received increasing attention in recent years[2] and is considered a relevant framework to promote healthy lifestyles through the promotion of physical activity (PA)[1,3,4] It highlights the necessary development of people's capacities through and for PA (e.g., confidence, motivation, social skills). Cairney et al. proposed a conceptual model linking PL, PA, and health throughout the life course and robust empirical evidence to support this model is beginning to emerge.[3] Indeed, the comprehensive benefits of PL and its underlying components have been well-documented, with studies highlighting the positive effects of physical, cognitive, affective, and social PL dimensions on PA level, well-being, and physical health, particularly for children.[5–8] However, despite this increasing interest and some encouraging results,[9] ongoing research is needed on proof of concept, understood here as evidence about the value of the concept, to support its viability before its implementation or further development. More evidence on PL is still needed, notably in terms of broadening the population of participants studied[10,11]

To explore the full potential of PL, its effects must be understood throughout the life course.[1,3,12] Each life event, and some life

* Corresponding author. 413, avenue Eugène Avinée 59120 Loos.
E-mail address: joseph.gandrieau@univ-lille.fr (J. Gandrieau).

periods, have a direct impact on an individual's PA levels.[13] "Emerging adulthood", defined as the time from the "end of adolescence" to "beginning of adulthood", which comes with responsibilities and important life changes, represents a challenging period for people to negotiate.[14] Indeed, new life events (e.g., leaving the family home, beginning careers, or enrolling in university) and a unique stage of development (e.g., maturational and developmental changes to brain and body, networking and social interactions through work and/or further study, expanded friendships and mutual support, family-oriented socialization, and learning about intimacy[15,16]) make this period a crucial. However, it is often overlooked in health lifestyle promotion,[17] making the analysis of PL during this period of major importance. To address this scientific challenge, the development of measurement tools is considered a crucial element.[10,11]

Early progress and uptake of holistic PL assessment tools have enriched the literature on PL measurement, leading to a raft of instruments, including the: PPLA-Q[18]; PLAY tools[19]; PL-C Quest[20]; CAEPL[21]; PFL[22]; Pre-PLAY[23]; PPLI[24]; and CAPL[25] Despite these methodological advances, research on PL assessment has largely focused on samples of school-age children, neglecting analysis of emerging adults as a population which is crucial to study, due to life changes described earlier. Recent research exposed the value of measuring PL in this population and outlined a preliminary tool available.[10,26] However, further evidence of the tool's validity is needed[10] since the tool was not initially constructed for use with a population other than young children.[26] An assessment tool's language, dimensions, domains, and complexity of items/tasks needs to reflect important specificities of a population under study. Among existing tools, only the PPLI was originally designed and validated for populations including emerging adults[24] Historically, the PPLI is one of the first tools to assess PL for physical education teachers and its validity has been explored in several populations, including seniors and adolescents in different cultures.[24,27–30] However, there is a need for a French version of the tool.

The lack of a PL assessment tool available in French prevents policymakers, professionals, and practitioners in French-speaking nations and communities from fully comprehending the importance of the issue in emerging adulthood. This period, which typically spans from 18 to 25 years old,[14] is crucial to an individual's PL journey as it transitions out of period which encompasses the last time they are required to participate in physical education in high school, when the most active individuals engage in organized PA. Emerging adulthood transits into important phases of beginning work and/or study, perhaps coinciding with starting a family, which all may severely impact PA levels. To promote the implementation and evaluation of French educational PL programs, and to support a thorough proof-of-concept analysis, a French-language PL measurement tool is needed.

Nevertheless, while expanding the research population on PL to include French emerging adults is an important challenge, there is also a sensitive issue to be considered: *is it really possible to measure PL?* This issue has been increasingly questioned by researchers[31–33] and to date, has remained a debated issue among PL experts. The monist philosophical foundation of this concept makes it challenging to quantify, as attempting to measure each dimension individually and then summing these scores, to calculate an overall PL value, may contradict its monist roots.[31,33] No studies have examined this issue and there is no empirical evidence to support the principle that "the sum of the parts (of the PL) is greater than the whole".[34] According to the "idealist" perspective[35] measuring PL is deemed pointless due to the uniqueness of this concept.[36] However, specific milestones can still be identified throughout the PL journey making it essential to provide a practical overview of PL during these critical periods.[10]

Undoubtedly, a PL measurement tool has the potential to make a significant contribution to both research on PL and the practices that support it. Before utilizing any measurement tool, it is imperative to rely on scientific evidence and ensure that the tool is both valid and reliable.[37] Validity and reliability are commonly measured through statistical methods, and it is crucial to evaluate these factors within the specific context in which the tool will be used.[37–39]

Therefore, the purpose of this study is to validate the French version of the PPLI for the emerging adult population. This validation process will involve evaluating the reliability, construct validity, and external validity of the tool, along with quantifying PA.

2. Methods

2.1. Participants

An initial questionnaire, in a digital format (Framasoft, Lyon, France), was distributed throughout a national university network in France. The recruitment process resulted in 2259 French emerging adults volunteering to participate in the study. Inclusion criteria included participants being enrolled at university from first to third year and to be aged from 18- to 26-years-old. Written informed consent was obtained from the participants before they entered the study. At the end of this first step, the participants were invited to continue being involved in the study by completing two other questionnaires: one survey immediately and another one two to three weeks later. In accordance with French law on data protection, an authorization to conduct this study was given by the Data Protection Officer of the University of Lille under number 202037 and the Ethics committee of the University of Lille (France). This authorization procedure secures data storage and guarantees the anonymity and withdrawal rights of participants.

2.2. Measurements

Three digital questionnaires were successively distributed to 2259 volunteer French university students between 8th February to 9th April 2021 (Fig. 1). To check its stability (test-retest reliability), the PPLI was completed a third time two to three weeks later.

The first questionnaire was composed of the PPLI[24] translated into French, along with other questions about individual characteristics (size, weight, sex, city of university and age). Initially PPLI was an 18-item test based on Likert scales (1–5) designed to address some key attributes of PL such as "*sense of self and self-confidence*", "*self-expression and communication with others*" and "*knowledge and understanding*".[24] The reliability level and internal validity of the PPLI were considered robust in previous studies[24,28,30,40] (Cronbach's α > 0.07; (r_w > 0.7; ICC>0.7). One crucial element during a translation process is making sure the item intent of the questionnaire and, therefore, the construct concepts are maintained across different languages and culture settings.[42] Accordingly, a back-translation technique[43] was used to translate the PPLI items from English to French. To maintain the linguistic equivalence of the scale, two bilingual PL experts who spoke both English and French, with knowledge of the PL concept, were engaged to independently translate the items (O.D; C.S). Their translations were then compared, and a triangulation process[44] was used to resolve any discrepancies between the two versions. The resulting French version was then submitted to two additional experts (T.D; F.P) to perform the reverse translation, from French to English. This additional step ensured that the translated items retained their original meaning and intention, and that there was no loss of meaning or distortion during the translation process.

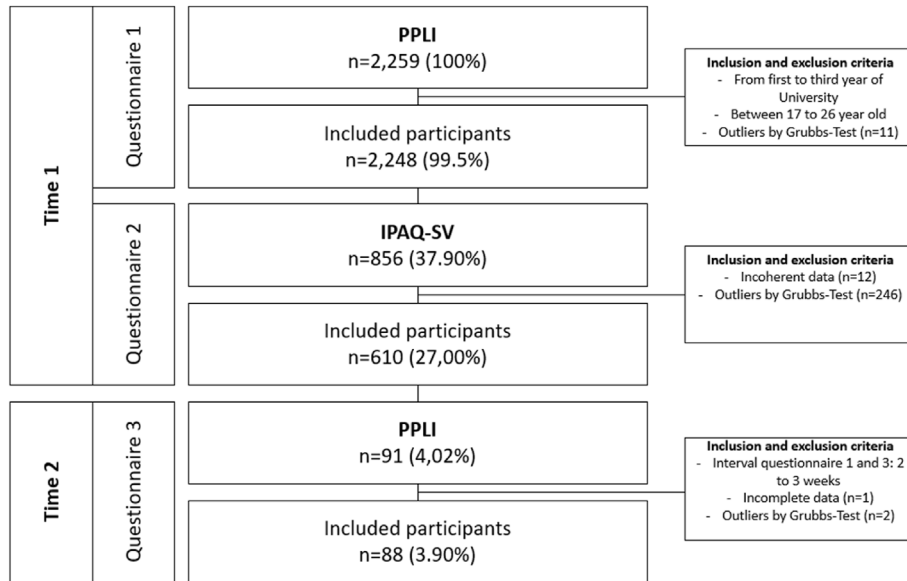The second instrument used was the International Physical

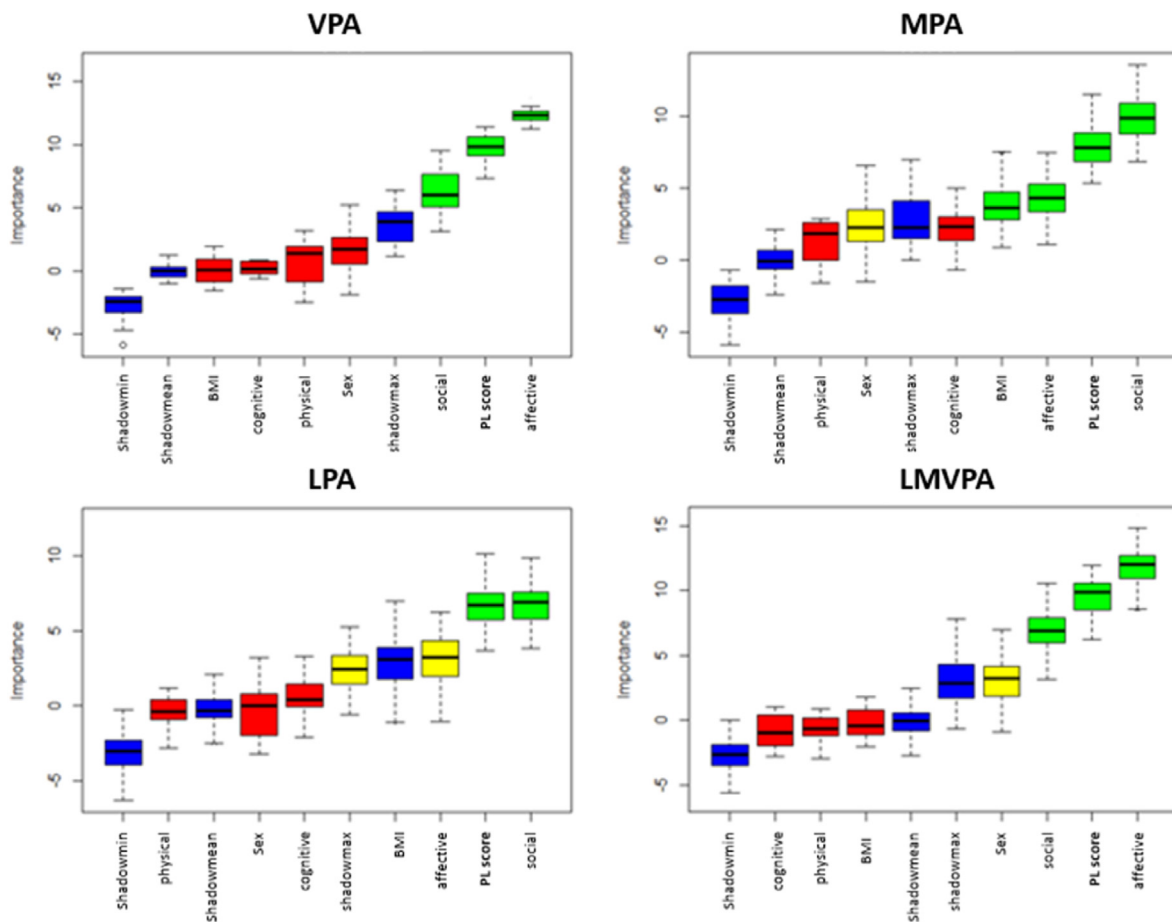**Fig. 1.** Flowchart of participant recruitment.

Fig. 2. Boruta results plot for VPA, MPA, LPA, and LMVPA data. Note: Blue boxplots correspond to minimal (shadowMin), average (shadowMean), and maximum (shadowMax) Z score of a shadow attribute. Red and green boxplots represent Z scores of respectively rejected and confirmed determinants. The yellow boxplots are tentative, that means the algorithm was not able to arrive to a conclusion about their importance.

Activity Questionnaire - Short Version (IPAQ-SV) translated in French. It has been translated into several languages and is the most

widely used questionnaire internationally.[45–47] This questionnaire retrospectively measured the overall amount of PA reported by

participants over the past 7 days, using 6 questions specifying three levels of intensity: vigorous PA (VPA), moderate PA (MPA) and light PA (LPA). There is a short (7 questions) and long (27 questions) version of this instrument. For this study we used the short version. Validation and reliability studies of the IPAQ short version in adults have shown that test-retest reliability is moderately high ($r_w = 0.74$), moderate for the criterion ($r_w = 0.41$) and moderately-high for concurrent validity ($r_w = 0.72$), indicating that the IPAQ short version is a valid and reliable tool for assessing physical activity levels in French-speaking adults.[47] In completing the questionnaire, the students were asked to exclude consideration of their formal, organized PA included in their studies, since our study does not focus on structured PA offered in the professional or school environment - which is too different, in France, due to the sports policies of each university.[48]

2.3. Statistical analysis

Exclusion criteria were based on the verification questions (n = 0) but also the analysis of incoherent data (e.g., reported PA participation being more than 7 days a week or more than 10,000 min per week, n = 12). Outliers were excluded using the statistical method of the Grubbs Test on each variable (n = 246). Details about selected samples are described in Table 1.

2.3.1. Internal validation and reliability
The sample was divided randomly into 2 balanced subsets (n = 1124 in each). An Exploratory Factor Analysis (EFA) of principal components with varimax rotation was conducted on the first subset to investigate the factor structure of the French PPLI. Outliers, Kayser-Meyer-Oklin (KMO) values and a Bartlett test were computed according to the recommendations of Broc et al.[49] The eigenvalue (>1) rule was used to determine the number of factors (4). Only items saturating a factor with a loading of >0.40 and a uniqueness value of <0.60 were retained.

Confirmatory Factor Analysis (CFA) was performed on the second subset to cross-validate and confirm the four-factor structure derived in the analysis. The goodness-of-fit test on the model was assessed using chi-square ($p > 0.05$), relative chi-square (χ^2/ddl; <0.05), root mean square error tests of approximation (RMSEA; <0.06), the Adjusted Goodness of Fit Statistic (AGFI; >0.90), Comparative Fit Index (CFI; >0.90), and Standardised Root Mean Square Residual Measures (SRMR; <0.08).[50,51] Nevertheless, since the RMSEA is sensitive to the kind of data reported in our study,[52] its interpretation remains flexible. To address the statistical limitations of Cronbach's α, which was initially used to check the internal consistency, Omega' ω[53,54] was also measured. To interpret the results, the same cutoffs as those used for Cronbach's α by Taber et al.[55] were employed. Alpha values were described as excellent (0.93—0.94), strong (0.91—0.93), reliable (0.84—0.90), robust (0.81), fairly high (0.76—0.95), high (0.73—0.95), good (0.71—0.91), relatively high (0.70—0.77), slightly low (0.68), reasonable (0.67—0.87), adequate (0.64—0.85), moderate (0.61—0.65), satisfactory (0.58—0.97), acceptable (0.45—0.98), sufficient (0.45—0.96), not satisfactory (0.4—0.55) and low (0.11). We measured two types of Omega: (a) McDonald's omega: this is the standard McDonald's

omega measure, which is based on the total variance of the measurement scale; (b) McDonald's hierarchical omega: this McDonald's omega measure takes into account the hierarchy of factors in a measurement scale. It is used when the items in a scale are organised into subscales or latent factors.

Intra-class correlation analysis (ICC) was performed in the context of reliability analysis (stability) with the data obtained from the first and third questionnaires by recording the intraclass correlation coefficient (ICC > 0.70[56]). Based on the recommendations of Koo and Li,[57] we performed Two-Way Random-Effects Models. To our knowledge, no study has examined the stability of the concept of PL and these data are important for choosing the test-retest time interval.[58] Consistent with interval chosen on specific studies about stability analysis of PL tools, we decided to adopt an interval of two to three weeks to examine tool stability.[59]

2.3.2. External validation: PL and PA association
The PL score was calculated following the recommendation of the PPLI method,[24] resulting in an overall score (n/45). Outliers (Grubbs test) and level of normality (Shapiro) were verified. The correlation between perceived PL and PA levels was computed by using the Spearman's rank correlation rhô due to the non-normal data distribution ($p < 0.05$). Correlation (rhô) was computed between the PL score and LPA, MPA, VPA and, the Light, Moderate, And Vigorous PA (LMVPA). We followed Cohen's recommendations[60] to consider the effect small when rhô = 0.1, medium when = 0.3, and large when = 0.5.

Feature selection by Boruta algorithm was computed to detect the strength of association between PL and PA. Sex, Body Mass Index (BMI = weight/size^2) and each independent PL dimension (physical, social, cognitive, and affective) were added to investigate the strength of the overall PL score as a determinant of PA. Boruta is a 'wrapper' approach developed around the random forest algorithm[61] used for feature selection. The Boruta algorithm calculates feature importance scores based on Z-scores of every input predictor concerning the shadow attribute randomly assigned.[62] A variable is considered relevant for classification if its feature importance score is greater than that of the best shadow feature (greater than "shadowmax"). Any variable that cannot satisfy this condition is reported as irrelevant (smaller than "shadowmax") and is discarded. Items with the highest feature importance scores (imp) are considered the best predictors of the dependent variable. The Boruta algorithm calculates average feature importance values based on 100 iterations of the random forest algorithm to increase the robustness of feature importance results.[62]

3. Results

3.1. Results of exploratory factor analysis (EFA) on subset 1

Results of the EFA on the French translation of PPLI showed that four factors were determined, based on parallel analysis scree plots (Table 2). The KMO index and the outcomes of the Bartlett's test were considered satisfactory (0.84, $p < 0.05$) and showed that a principal components analysis (PCA) could be computed.[49] Ten out of 18 items were deleted after initial factor analysis resulted in the design of a new survey of 8 items (Table 2 and annexe1) PCA resulted in a structure of 4 factors explaining 51% of the variance. For each factor, Cronbach's α and omega'ω scores were 'relatively high' (>0.7) to 'reliable' (0.84—0.90[55]). Cronbach's α and hierarchical omega' ω scores for the entire set of the French PPLI confirms the overall internal consistency of the assessment tool composed of four dimensions (0.88; 77).

Table 1
Characteristics of the sample of participants.

	Questionnaire 1	Questionnaire 2	Questionnaire 3
Whole sample	2248	856	88
Women	1015	546	54
Men	1233	310	34
Age (mean ± sd)	19 ± 1.53	19 ± 1.04	19 ± 1.16

Table 2
Exploratory factor analysis of the French PPLI.

Item	Factor 1	Factor 2	Factor 3	Factor 4
I'm able to apply learnt motor skills to other physical activities	–	–	0.67	–
I have positive attitude and interest in sports	–	–	–	0.62
I appreciate myself or others doing sport	–	–	–	0.62
I am able to apply PE knowledge in the long run	–	–	0.58	–
I possess self-management skills for fitness	–	0.69	–	–
I possess self-evaluation skills for health	–	0.65	–	–
I have strong communication skills	0.86	–	–	–
I have strong social skills	0.86	–	–	
Factor variance	19	12	10	10
Cumulative of variance explained (%)	19	31	41	51
Cronbach α	0.88	0.75	0.75	0.89
Omega ω	0.88	0.75	0.75	0.89
ICC [IC95%]	0.86	0.95	0.83	0.86
	[0.78; 0.90]	[0.92; 0.97]	[0.75; 0.89]	[0.79; 0.91]

Note. The ICC for the full test (8 items) is 0.91[0.87; 0.95] and the Cronbach α and hierarchical Omega ω is respectively 0.88 and 0.77.

3.2. Results of Confirmatory Factor Analysis (CFA) on subset 2

Based on the EFA results, CFA was computed to confirm this four-factor structure. Factor validity was satisfactory due to factor loading of all items above the standard of 0.40 with a minimum value of 0.58 (item 6). The goodness-of-fit index was satisfactory with χ^2/ddl <0.05, RMSEA <0.06, AGFI >0.90, SRMR <0.08, CFI >0.90. The reliability (stability) values within two to three weeks of the four-factor test were good to excellent, ranging from 0.83 to 0.95 for each factor and 0.91 for the total score. Based on this new version, the tools revealed high values of ICC for all dimension respectively 0.86, 0.95, 0.83, 0.86. Finally, ICC was 0.91 for full test.

3.3. Results of the correlation between PL and PA

The PL median score of participants was 32.5/45. The reported median of LMVPA is 2111.5 mets.week^1 which allows it to be included in our sample in the 2nd level of the IPAQ-SV guidelines (600 < LMVPA <3000 mets.week^{-1}) (Table 3). Table 4 displays the correlation between PL and PA intensities levels. The scores of PL were significantly correlated with each level of PA intensity ($p < 0.01$) Each correlation test reached statistical significance levels, but with small to medium values. The highest correlation value was shown for the LMVPA (0.33, $p < 0.01$). In contrast, the smallest correlation values were revealed between PL and LPA, but were still statistically significant (0.18, $p < 0.01$).

Table 3
Descriptive results about LMVPA, VPA, MPA, LPA and PL scores (n = 610).

	Median	Q1	Q3	Skew	Kurtosis
PL Score (/45)	32.5	30	36	−0.65	−0.17
VPA	1319.8	400	2400	−0.03	0.86
MPA	480	240	960	1.08	0.54
LPA	292.6	99	495	1.06	0.58
LMVPA	2111.5	1026.0	2111.5	0.12	0.77

Note. The VPA, MPA, LPA, and LMVPA score is given in Mets.minutes.week^{-1}.

Table 4
Results of correlation (rhô) between PL and VPA, MPA, LPA, and LMVPA.

	LMVPA	VPA	MPA	LPA
PL	0.33[a]	0.31[a]	0.21[a]	0.18[a]

Note: Power>0.98.
[a] Correlation is significant at the 0.01 level.

3.4. Results of Boruta algorithm

The feature importance scores generated by the Boruta algorithm for each PA level of intensity are reported in Table 5 and Fig. 2. Boruta demonstrated that, overall, the PL score is confirmed as an important predictor of PA for VPA, MPA, LPA, and LMVPA. Nevertheless, the median feature importance score suggested that the overall PL score is not the most important predictor of PA participation, among other variables like Sex, BMI, social, affective, physical, and cognitive dimension (Table 5, Fig. 1). The overall PL score has an importance slightly below the most impactful determinant of each intensity − the affective score for the VPA and LPA and social score for the MPA and LMVPA.

4. Discussion

The aim of this study was to compute the validation process of the French PPLI with a sample of emerging adults. Results showed the French PPLI was a valid and reliable tool to measure PL, constituted by four dimensions - social, affective, physical, and cognitive − representing together a higher order structure. Based on correlation analysis and use of the Boruta algorithm, this study provides additional evidence of the value of promoting PL to facilitate PA engagement in this population.

4.1. A reliable and valid four-dimension-tool

Our results exposed great internal validity and acceptable reliability for the French PPLI in the sample of emerging adults. The first factor is composed of items 10 and 11 and related to the domain "*self-expression and communication with others*" of Sum et al.[24] (a social dimension). The second factor was composed of items 7 and 8, relating to skills and knowledge for managing one's health. It was, therefore, closer to a domain of "*knowledge and know-how for one's PA and health*" than "*knowledge and understanding*" of Sum et al.[24] (cognitive dimension). The third factor was constituted by items 3 and 6 which relate to "*control of the environment*" (physical dimension). Finally, the fourth factor is composed of items 4 and 5, relating to "*positive affect in practice*" (affective dimension). The levels of internal consistency of each factor were "relatively high" to "reliable" [55] with values ranging from 0.75 to 0.89 (α; ω) and the test-retest reliability values were largely acceptable. Thus, the French PPLI was composed of 4 dimensions - physical, cognitive, affective, and social - and converged with findings on the most recent conceptual frameworks of PL[12,63] including a specific study on PL emerging adults.[10]

Table 5
Variable importance information obtained with Boruta algorithm.

		Mean Imp	Median Imp	Min Imp	Max Imp	Decision
VPA	Sex	1.72	1.77	−1.87	5.26	Rejected
	BMI	0.02	0.08	−1.55	1.95	Rejected
	Affective	12.57	12.37	11.23	16.39	**Confirmed**
	Social	6.39	6.03	3.15	9.52	**Confirmed**
	Physical	0.84	1.41	−2.47	3.25	Rejected
	Cognitive	0.39	0.17	−0.62	2.66	Rejected
	PL score	9.87	9.89	7.39	1.95	**Confirmed**
MPA	Sex	2.27	2.30	−1.91	6.60	Rejected
	BMI	3.78	3.66	−0.66	7.61	**Confirmed**
	Affective	4.44	4.37	1.12	8.60	**Confirmed**
	Social	9.90	9.90	6.88	13.63	**Confirmed**
	Physical	1.31	1.84	−1.57	2.92	Rejected
	Cognitive	2.14	2.35	−1.31	5.90	Rejected
	PL score	7.92	7.82	5.39	11.55	**Confirmed**
LPA	Sex	−0.38	−0.41	−2.81	2.40	Rejected
	BMI	−0.19	−0.01	−3.25	3.21	Rejected
	Affective	6.69	6.71	3.71	10.16	**Confirmed**
	Social	0.66	0.41	−2.13	3.31	**Confirmed**
	Physical	3.18	3.22	−1.04	6.26	Rejected
	Cognitive	2.48	2.45	−0.57	5.23	Rejected
	PL score	6.74	6.89	2.47	12.41	**Confirmed**
LMVPA	Sex	3.14	3.24	−2.60	9.24	Rejected
	BMI	−0.16	−0.44	−2.05	1.80	Rejected
	Affective	11.87	11.99	8.25	15.55	**Confirmed**
	Social	6.94	6.89	3.18	10.56	**Confirmed**
	Physical	−0.90	−0.67	−3.50	0.88	Rejected
	Cognitive	−0.92	−0.97	−2.83	0.99	Rejected
	PL score	9.49	9.90	6.16	11.95	**Confirmed**

The quality of a tool can be also assessed through its pragmatic nature. With 8 items, this new tool has a real pragmatic quality (based on speed of administration) to assess the level of PL in four dimensions. Nevertheless, with two items per dimension, the tool does not comply with the recommendations stating that below 3 items per factor there may be a lack of reliability.[64–66] The pragmatic aspect of the tool hinders its scientific robustness. Nevertheless, the low number of items does not seem to be detrimental to the stability of the tool, which is considered a crucial element of a measurement tool.[58] Moreover, some studies point out that it may be acceptable to use tools with fewer than three items.[67,68]

The retained items and dimensions were different from those in previous tool versions,[24,29,41] for example, revealing four constitutive factors, instead of three, resulting from the first analysis dimension computed by Sum et al.[24] These differences in dimensions are not surprising because each of the contexts differed in which the PPLI has been analyzed.[24,27,29,41] These results may be explained by the culturally embedded nature of the PL concept,[1,4] and highlight the importance of analyzing in detail the PL structure concept for each population studied. Nevertheless, our results continue to provide some evidence on the holistic nature of PL.[1]

The acceptable hierarchical omega index ω indicates the presence of a superior construct underlying the four distinct dimensions. This index is commonly used to evaluate the internal consistency of a multilateral hierarchical model with subscales or factors. A high omega-h value suggests that scale scores result from the largely dependent contributions of unique facets, with a significant contribution from a latent construct common to all facets.[69] In line with the theoretical foundation of the PL concept, our hierarchical omega index ω suggests that the tool measures an underlying construct higher than the four distinct dimensions. Our results are consistent with recent empirical studies that support the construct validity of the PL concept as a superior construct to the simple juxtaposition of distinct dimensions.[70,71]

4.2. Association between PL and PA

Although recent studies have provided some evidence for the effectiveness of PL in promoting PA, it is still not enough for PL to be widely accepted "*as best practice in reduction of non-communicable disease or the promotion of physical activity participation*"[72] (p14). Our study contributes to this ongoing issue by providing new evidence, but it is important to note that proof of concept research is still in its infancy.

Our correlation results demonstrated a significant association between PL and the four PA intensities, and the implementation of the Boruta algorithm confirms that PL can be considered an important determinant of each PA intensities, including vigorous, moderate, light PA. This finding is consistent with data reported in previous studies that have also found a positive relationship between PL and PA across different intensity levels.[5,6,23,73,74] Moreover, our Boruta results suggest that PL is particularly associated with vigorous levels of PA, which appears to be the most intensity recommended for health benefits.[75] However, the positive association between PL and all levels of PA intensity is consistent with the results of Choi et al.[73] who found statistically significant associations across all intensity levels. Therefore, although PL may be more effective in promoting vigorous PA during emerging adulthood, it is still a relevant and valuable concept for promoting PA at all intensity levels which is crucial, given the diverse motives and participation patterns in PA among individuals[76]

The strength of these association was in line with data reported in previous studies[73,74] which found statistically significant, but weaker, associations ($r^2 = 0.10$ to 0.35). This small to medium association value raised questions about the importance of PL as "the cornerstone" of a physically active lifestyle, among other PA determinants (e.g., territorial organization, socio-professional category, peer, and family support) and specifically individuals' PA determinants (i.e., motor competences, PA enjoyment, or motivation). These findings encourage future studies to compare the importance of a PL score with other determinants of the PA socio-

ecological model.[77] Considering the accepted strong theoretical value of PL in explaining PA,[1,4] it would be crucial to question our ability to actually capture the value of the concept.

Our Boruta results also prompted us to consider a question: was the weak to moderate association between PL and PA attributable to the tools used or to current theorising behind the concept itself? Our results contradict the claims of PL scholars, as we found that the overall PL score did not have the most significant impact on explaining engagement in PA, considering its constitutive dimensions including social, affective, cognitive, and physical factors. The social (MPA, LPA) and affective (VPA, LMVPA) dimensions were more important than the sum of the four dimensions to explain PA involvement. Here, the superior construct that PL represents[70] - going beyond the juxtaposition of its constituent elements[34] - does not really expose its added value over these separate dimensions to explain PA. Our hypothesis is rather that further studies should examine an "evaluation effect" before questioning the concept itself. Indeed, some authors have raised questions over the weakness of linear and more simple arithmetic approaches, consisting of summative scores by dimension,[31] highlighting the need for a more comprehensive and specific tool for the study population scrutinized.[10] The weak to moderate association observed between overall PL and PA might be greater once this challenge - to capture this complex concept - may be overcome. Future studies are needed to investigate associations with other assessment tools and different ways of scoring, considering more complex methods and the interdependence of constituent dimensions.

4.3. Limits and perspectives

In this study, we aimed to provide evidence on the validity and reliability of the French PPLI for emerging adults. PPLI might help teachers to guide and support students in their PL journeys by analyzing student PL levels. PPLI will also be useful to study the importance of PL and how it changes during the key period of emerging adulthood. However, additional study is still needed to provide further support for the value of this tool. From a methodological point of view, the population was not recruited with a random selection of universities. Moreover, emerging adults were only recruited here from a student population and thus may represent a bias in the representativeness of this sample for the French emerging adult population. External validity could also be improved by not only focusing on the level of PA at the immediate moment, but by focusing on a longitudinal follow-up of PA levels. Researchers are invited to use more objective types of PA measurements[78] coupled with more subjective tools (e.g., daily digital diary,[79] Indeed, the self-reported questionnaire has some limitations, particularly about accidental PA — referring to unplanned or unintentional physical activity resulting from everyday actions such as climbing stairs, gardening, cleaning the house or walking to do errands.

5. Conclusion

Our results indicated that the PPLI is a reliable and valid instrument to measure PL levels in French emerging adults. The French PPLI is a tool that can be used by teachers and PA practitioners to evaluate the PL development of French emerging adults. External validation measures converged with those reported in previous studies by displaying a significant, small to medium PL association with different PA intensities, underscoring the validity of this tool. It supports the importance of developing PL for all periods of life, other than childhood only. The findings point to the need for future studies on the analysis of the relations between PA and PL during this crucial period, but questions about the value of the concept, and how to measure it, remain requiring further study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesf.2023.06.001.

References

1. Whitehead M. *Physical Literacy: Throughout the Lifecourse*. Routledge; 2010. https://doi.org/10.4324/9780203881903.
2. Young L, O'Connor J, Alfrey L. Physical literacy: a concept analysis. *Sport Educ Soc*. 2019;25(8):946–959. https://doi.org/10.1080/13573322.2019.1677586.
3. Cairney J, Dudley D, Kwan M, Bulten R, Kriellaars D. Physical literacy, physical activity and health: toward an evidence-informed conceptual model. *Sports Med*. 2019;49(3):371–383. https://doi.org/10.1007/s40279-019-01063-3.
4. Whitehead M. *Physical Literacy across the World*. Routledge; 2019. https://doi.org/10.4324/9780203702697.
5. Brown D, Dudley DA, Cairney J. Physical literacy profiles are associated with differences in children's physical activity participation: a latent profile analysis approach. *J Sci Med Sport*. 2020. https://doi.org/10.1016/j.jsams.2020.05.007. S1440244019317116.
6. Caldwell HAT, Di Cristofaro NA, Cairney J, Bray SR, MacDonald MJ, Timmons BW. Physical literacy, physical activity, and health indicators in school-age children. *Int J Environ Res Publ Health*. 2020;17(15):5367. https://doi.org/10.3390/ijerph17155367.
7. Clark HJ, Dudley D, Barratt J, Cairney J. Physical literacy predicts the physical activity and sedentary behaviours of youth. *J Sci Med Sport*. 2022;25(9): 750–754. https://doi.org/10.1016/j.jsams.2022.04.008.
8. Melby PS, Nielsen G, Brønd JC, Tremblay MS, Bentsen P, Elsborg P. Associations between children's physical literacy and well-being: is physical activity a mediator? *BMC Publ Health*. 2022;22(1):1267. https://doi.org/10.1186/s12889-022-13517-x.
9. Kwan MYW, Graham JD, Bedard C, Bremer E, Healey C, Cairney J. Examining the effectiveness of a pilot physical literacy–based intervention targeting first-year university students: the PLUS program. *Sage Open*. 2019;9(2), 215824401985024. https://doi.org/10.1177/2158244019850248.
10. Gandrieau J, Schnitzler C, Cairney J, et al. Development of ELIP to assess physical literacy for emerging adults: a methodological and epistemological challenge. *Res Q Exerc Sport*. 2023:1–14. https://doi.org/10.1080/02701367.2022.2125927.
11. Tremblay MS, Lloyd M. Physical literacy measurement - the missing piece. *Phys Health Educ Can J*. 2010;76:26–30.
12. Keegan RJ, Barnett LM, Dudley DA, et al. Defining physical literacy for application in Australia: a modified delphi method. *J Teach Phys Educ*. 2019;38(2): 105–118. https://doi.org/10.1123/jtpe.2018-0264.
13. Gropper H, John JM, Sudeck G, Thiel A. The impact of life events and transitions on physical activity: a scoping review. In: Prince Ware S, ed. *PLOS ONE*. 2020, e0234794. https://doi.org/10.1371/journal.pone.0234794.
14. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. *Am Psychol*. 2000;55(5):469–480.
15. Berndt TJ, Savin-Williams RC. Peer relations and friendships. In: *Handbook of Clinical Research and Practice with Adolescents. Wiley Series on Personality Processes*. John Wiley & Sons; 1993:203–219.
16. Hochberg Z, Konner M. Emerging adulthood, a pre-adult life-history stage. *Front Endocrinol*. 2020;10:918. https://doi.org/10.3389/fendo.2019.00918.
17. Howells K, Coppinger T. The forgotten age phase of healthy lifestyle promotion? A preliminary study to examine the potential call for targeted physical activity and nutrition education for older adolescents. *Int J Environ Res Publ Health*. 2022;19(10):5970. https://doi.org/10.3390/ijerph19105970.

18. Mota J, Martins J, Onofre M. Portuguese Physical Literacy Assessment Questionnaire (PPLA-Q) for adolescents (15—18 years) from grades 10—12: development, content validation and pilot testing. *BMC Publ Health*. 2021;21(1): 2183. https://doi.org/10.1186/s12889-021-12230-5.

19. Caldwell HA, Di Cristofaro NA, Cairney J, Bray SR, Timmons BW. Measurement properties of the physical literacy assessment for youth (PLAY) tools. *Appl Physiol Nutr Metabol*. 2021;46(6):571—578. https://doi.org/10.1139/apnm-2020-0648.

20. Barnett LM, Mazzoli E, Hawkins M, et al. Development of a self-report scale to assess children's perceived physical literacy. *Phys Educ Sport Pedagog*. 2020: 1—26. https://doi.org/10.1080/17408989.2020.1849596. Published online December 22.

21. Chen ST, Tang Y, Chen PJ, Liu Y. The development of Chinese assessment and evaluation of physical literacy (CAEPL): a study using delphi method. *Int J Environ Res Publ Health*. 2020;17(8):2720. https://doi.org/10.3390/ijerph17082720.

22. Lodewyk KR. Early validation evidence of the Canadian practitioner-based assessment of physical literacy in secondary physical education. *Phys Educ*. 2019;76(3):634—660.

23. Cairney J, Clark HJ, James ME, Mitchell D, Dudley DA, Kriellaars D. The preschool physical literacy assessment tool: testing a new physical literacy tool for the early years. *Front Pediatr*. 2018;6:138. https://doi.org/10.3389/fped.2018.00138.

24. Sum RKW, Ha ASC, Cheng CF, et al. In: Ginsberg SD, ed. *Construction and Validation of a Perceived Physical Literacy Instrument for Physical Education Teachers*. PLOS ONE; 2016, e0155610. https://doi.org/10.1371/journal.pone.0155610, 11(5).

25. Longmuir PE, Boyer C, Lloyd M, et al. The Canadian Assessment of Physical Literacy: methods for children in grades 4 to 6 (8 to 12 years). *BMC Publ Health*. 2015;15(1):767. https://doi.org/10.1186/s12889-015-2106-6.

26. Kleis RR, Dlugonski D, Baker CS, Hoch JM, Hoch MC. Examining physical literacy in young adults: psychometric properties of the PLAYself. *Appl Physiol Nutr Metabol*. 2022;47(9):926—932. https://doi.org/10.1139/apnm-2022-0062.

27. Liu CY, Lin LLC, Sheu JJ, Sum RKW. Psychometric validation of senior perceived physical literacy instrument. *Int J Environ Res Publ Health*. 2022;19(11):6726. https://doi.org/10.3390/ijerph19116726.

28. Ma RS, Sum RKW, Hu YN, Gao TY. Assessing factor structure of the simplified Chinese version of perceived physical literacy instrument for undergraduates in mainland China. *J Exerc Sci Fit*. 2020;18(2):68—73. https://doi.org/10.1016/j.jesf.2020.01.001.

29. Sum RKW, Cheng CF, Wallhead T, Kuo CC, Wang FJ, Choi SM. Perceived physical literacy instrument for adolescents: a further validation of PPLI. *J Exerc Sci Fit*. 2018;16(1):26—31. https://doi.org/10.1016/j.jesf.2018.03.002.

30. López-Gil JF, Martínez-Vizcaíno V, Tárraga-López PJ, García-Hermoso A. Cross-cultural adaptation, reliability, and validation of the Spanish perceived physical literacy instrument for adolescents (S-PPLI). *J Exerc Sci Fit*. 2023;21(3): 246—252. https://doi.org/10.1016/j.jesf.2023.03.002.

31. Chen A. A clash of fundamental assumptions: can/should we measure physical literacy? *J Sport Health Sci*. 2020;9(2):149—151. https://doi.org/10.1016/j.jshs.2019.11.002.

32. Edwards LC, Bryant AS, Keegan RJ, Morgan K, Cooper SM, Jones AM. 'Measuring' physical literacy and related constructs: a systematic review of empirical findings. *Sports Med*. 2018;48(3):659—682. https://doi.org/10.1007/s40279-017-0817-9.

33. Green NR, Roberts WM, Sheehan D, Keegan RJ. Charting physical literacy journeys within physical education settings. *J Teach Phys Educ*. 2018;37(3): 272—279. https://doi.org/10.1123/jtpe.2018-0129.

34. Dudley D. Physical literacy: when the sum of the parts is greater than the whole. *J Phys Educ Recreat Dance*. 2018;89(3):7—8. https://doi.org/10.1080/07303084.2018.1418998.

35. Edwards LC, Bryant AS, Keegan RJ, Morgan K, Jones AM. Definitions, foundations and associations of physical literacy: a systematic review. *Sports Med*. 2017;47(1):113—126. https://doi.org/10.1007/s40279-016-0560-7.

36. Green NR, Roberts WM, Sheehan D, Keegan RJ. Charting physical literacy journeys within physical education settings. *J Teach Phys Educ*. 2018;37(3): 272—279. https://doi.org/10.1123/jtpe.2018-0129.

37. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539—549. https://doi.org/10.1007/s11136-010-9606-8.

38. Gunnell KE, Schellenberg BJI, Wilson PM, Crocker PRE, Mack DE, Zumbo BD. A review of validity evidence presented in the journal of sport and exercise psychology (2002—2012): misconceptions and recommendations for validation research. In: Zumbo BD, Chan EKH, eds. *Validity And Validation In Social, Behavioral, and Health Sciences. Vol 54. Social Indicators Research Series*. Springer International Publishing; 2014:137—156. https://doi.org/10.1007/978-3-319-07794-9_8.

39. Messick S. Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Am Psychol*. 1995;50(9):741—749. https://doi.org/10.1037/0003-066X.50.9.741.

40. Yilmaz A, Kabak S. Perceived physical literacy scale for adolescents (PPLSA): validity and reliability study. *Int J Educ Literacy Stud*. 2021;9(1):159—171.

41. Ma RS, Sum RKW, Hu YN, Gao TY. Assessing factor structure of the simplified Chinese version of perceived physical literacy instrument for undergraduates in mainland China. *J Exerc Sci Fit*. 2020;18(2):68—73. https://doi.org/10.1016/

42. Hawkins M, Cheng C, Elsworth GR, Osborne RH. Translation method is validity evidence for construct equivalence: analysis of secondary data routinely collected during translations of the Health Literacy Questionnaire (HLQ). *BMC Med Res Methodol*. 2020;20(1):130. https://doi.org/10.1186/s12874-020-00962-8.

43. Brislin RW. Back-translation for cross-cultural research. *J Cross Cult Psychol*. 1970;1(3):185—216. https://doi.org/10.1177/135910457000100301.

44. Lincoln YS, Guba EG. *Naturalistic Inquiry*. SAGE; 1985.

45. Craig CL, Marshall AL, Sj??Str??M M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381—1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB.

46. Lee PH, Macfarlane DJ, Lam T, Stewart SM. Validity of the international physical activity questionnaire short form (IPAQ-SF): a systematic review. *Int J Behav Nutr Phys Activ*. 2011;8(1):115. https://doi.org/10.1186/1479-5868-8-115.

47. Sember V, Meh K, Sorić M, Starc G, Rocha P, Jurak G. Validity and reliability of international physical activity questionnaires for adults across EU countries: systematic review and meta analysis. *Int J Environ Res Publ Health*. 2020;17(19): 7161. https://doi.org/10.3390/ijerph17197161.

48. Dec M, Larras B, Patoux L, Genin P, Bruchet L. *Pratique d'activités physiques et sportives et sédentarité chez les étudiants en formations universitaires en France — Résultats d'une enquête nationale*. ONAPS; 2023:34. https://onaps.fr/wp-content/uploads/2023/01/2023-01-17-Rapport-final-enque%CC%82te-e%CC%81tudiante-VF-logossedito.pdf. Accessed June 3, 2023.

49. Broc G, Carlsberg M, Cazauvieilh C, Faury S, Loyal D, Atzeni T. *Stats Faciles Avec R: Guide Pratique*. De boeck supérieur; 2016.

50. Hooper D, Coughlan J, Mullen MR. *Structural Equation Modelling: Guidelines for Determining Model Fit*. 2008;1(6):53—60.

51. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J*. 1999;6(1):1—55. https://doi.org/10.1080/10705519909540118.

52. Shi D, Maydeu-Olivares A, Rosseel Y. Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA. *Struct Equ Model Multidiscip J*. 2020;27(1):1—15. https://doi.org/10.1080/10705511.2019.1611434.

53. McDonald RP. The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *Br J Math Stat Psychol*. 1970;23(1): 1—21. https://doi.org/10.1111/j.2044-8317.1970.tb00432.x.

54. Béland S, Cousineau D, Loye N. Utiliser le coefficient omega de McDonald à la place de l'alpha de Cronbach. *McGill J Educ*. 2018;52(3):791—804. https://doi.org/10.7202/1050915ar.

55. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ*. 2018;48(6):1273—1296. https://doi.org/10.1007/s11165-016-9602-2.

56. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420—428. https://doi.org/10.1037/0033-2909.86.2.420.

57. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155—163. https://doi.org/10.1016/j.jcm.2016.02.012.

58. Polit DF. Getting serious about test—retest reliability: a critique of retest research and some recommendations. *Qual Life Res*. 2014;23(6):1713—1720. https://doi.org/10.1007/s11136-014-0632-9.

59. Barnett LM, Mazzoli E, Bowe SJ, Lander N, Salmon J. Reliability and validity of the PL-C Quest, a scale designed to assess children's self-reported physical literacy. *Psychol Sport Exerc*. 2022;60, 102164. https://doi.org/10.1016/j.psychsport.2022.102164.

60. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum; 1988.

61. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5—32. https://doi.org/10.1023/A:1010933404324.

62. Kursa MB, Rudnicki WR. Feature selection with the **Boruta** package. *J Stat Software*. 2010;36(11). https://doi.org/10.18637/jss.v036.i11.

63. Martins J, Onofre M, Mota J, et al. International approaches to the definition, philosophical tenets, and core elements of physical literacy: a scoping review. *Prospects*. 2020. https://doi.org/10.1007/s11125-020-09466-1. Published online May 28.

64. DeVellis RF, Thorpe CT. *Scale Development: Theory and Applications*. fifth ed. SAGE Publications, Inc; 2022.

65. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychol Methods*. 1999;4(1):84—99. https://doi.org/10.1037/1082-989X.4.1.84.

66. Raubenheimer J. An item selection procedure to maximise scale reliability and validity. *SA J Ind Psychol*. 2004;30(4). https://doi.org/10.4102/sajip.v30i4.168.

67. Proctor SL, Hoffmann NG, Kopak AM. An ultra-brief 2-item depression screening tool for correctional populations. *J Correct Health Care*. 2021;27(1): 36—39. https://doi.org/10.1089/jchc.19.06.0049.

68. Wanous JP, Reichers AE, Hudy MJ. Overall job satisfaction: how good are single-item measures? *J Appl Psychol*. 1997;82(2):247—252. https://doi.org/10.1037/0021-9010.82.2.247.

69. Widhiarso W, Ravand H. Estimating reliability coefficient for multidimensional measures: a pedagogical illustration. *Rev Psychol*. 2014;21:111—121.

70. Cairney J, Clark H, Dudley D, Kriellaars D. Physical literacy in children and youth—a construct validation study. *J Teach Phys Educ*. 2019;38(2):84—90. https://doi.org/10.1123/jtpe.2018-0270.

71. Britton Ú, Belton S, Peers C, et al. Physical literacy in children: exploring the

construct validity of a multidimensional physical literacy construct. *Eur Phys Educ Rev*. 2023;29(2):183–198. https://doi.org/10.1177/1356336X221131272.

72. Dudley D, Cairney J, Wainwright N, Kriellaars D, Mitchell D. Critical considerations for physical literacy policy in public health, recreation, sport, and education agencies. *Quest*. 2017;69(4):436–452. https://doi.org/10.1080/00336297.2016.1268967.

73. Choi SM, Sum RKW, Leung EFL, Ng RSK. Relationship between perceived physical literacy and physical activity levels among Hong Kong adolescents. *PLoS One*. 2018;13(8). https://doi.org/10.1371/journal.pone.0203105.

74. Ma RS, Sum RKW, Li MH, Huang Y, Niu XL. Association between physical literacy and physical activity: a multilevel analysis study among Chinese undergraduates. *Int J Environ Res Publ Health*. 2020;17(21):7874. https://doi.org/10.3390/ijerph17217874.

75. Hills AP, Dengel DR, Lubans DR. Supporting public health priorities: recommendations for physical education and physical activity promotion in schools.

Prog Cardiovasc Dis. 2015;57(4):368–374. https://doi.org/10.1016/j.pcad.2014.09.010.

76. de Hollander EL, Proper KI. Physical activity levels of adults with various physical disabilities. *Prev Med Rep*. 2018;10:370–376. https://doi.org/10.1016/j.pmedr.2018.04.017.

77. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJ, Martin BW. Correlates of physical activity: why are some people physically active and others not? *Lancet*. 2012;380(9838):258–271. https://doi.org/10.1016/S0140-6736(12)60735-1.

78. Vanhelst J. Quantification de l'activité physique par l'accélérométrie. *Rev DÉpidémiologie Santé Publique*. 2019;67(2):126–134. https://doi.org/10.1016/j.respe.2018.10.010.

79. Derigny T, Schnitzler C, Remmers T, et al. Catch me if you can! How French adolescents seize social occasions and opportunities to be active. *BMC Publ Health*. 2022;22(1):1332. https://doi.org/10.1186/s12889-022-13746-0.