Methods and Applications

Development and Testing of Physical Literacy Scales for Chinese Elementary School Students — China, 2022

Mengyu Li[1]; Ziqing Zan[1]; Lizhu Liu[1]; Hanxin Gu[1]; Yu Zhang[1]; Lili You[1,#]

ABSTRACT

Background: This study aimed to develop and validate test-based physical literacy scales (PLSs) for primary school students across different grades.

Methods: Data were collected through a field survey conducted from June 1 to July 31, 2022, involving 3,275 primary school students from four provinces in China. The questionnaires assessed four domains: physical knowledge, physical competence, physical motivation, and physical behavior. The Rasch model was employed for psychometric analysis.

Results: The variance explained by measures for the PLSs was 53.1%, 50.3%, and 54.7%, all exceeding the 50% threshold, confirming unidimensionality and robust internal consistency. This enabled effective differentiation among students with varying proficiency levels. The item-person map demonstrated optimal alignment between item difficulty and participant ability levels. Most items showed favorable fit statistics, with Infit mean square (MNSQ) and Outfit MNSQ values ranging between 0.5 and 1.5.

Conclusions: The PLSs demonstrate validity and reliability in measuring physical literacy among Chinese primary school students across four components: physical knowledge, physical competence, physical motivation, and physical behavior. The scales measure a unidimensional construct, supporting the use of summed total scores for assessment.

Physical literacy (PL) has evolved from the concept of health literacy (HL). The fundamental components of HL comprise emotional dimensions (motivation and confidence), physical dimensions (physical capability), and cognitive dimensions (knowledge and understanding) (*1*). PL represents a comprehensive construct that encompasses motivation, self-confidence, physical competence, motor skill execution, and active engagement in physical activities (*2*–*4*). Recent researches has extensively documented the benefits of physical activities, particularly those of moderate to vigorous intensity, for children and adolescents, including improvements in cardiovascular health, visual acuity, and mental well-being (*5*–*7*).

Despite these recognized benefits, the current state of physical activity among primary school students in China remains concerning, characterized by insufficient understanding of physical exercise and declining physical fitness (*8*). PL enables primary school students to develop fundamental comprehension of exercise and health, thereby fostering health-promoting behaviors and competencies (*9*). Students with strong PL competencies consistently demonstrate superior performance in physical activities. Therefore, PL assessment serves as a crucial tool for monitoring and evaluating the effectiveness of physical education among primary school students, ultimately informing and enhancing national-level policies and interventions.

Current research and assessment tool development for PL in China have predominantly focused on adolescent populations, particularly middle and high school students. A significant gap exists in the availability of standardized measurement tools for assessing PL among primary school students in China. This study addresses this gap by developing a comprehensive questionnaire specifically designed to measure PL in primary school students, taking into account the developmental characteristics of students across different grade levels.

METHODS

Study Population

This study employed a multistage cluster sampling approach conducted from June 1 to July 31, 2022. Three provinces were strategically selected to represent the eastern, central, and western regions of China, with one city randomly chosen from each province.

Additionally, a highly economically developed municipality in southern China was included to enhance regional representation. Within each selected city, we randomly selected one urban and one rural primary school. From each school, 1–2 classes were randomly sampled across three grade level groups (grades 1–2, 3–4, and 5–6). All students within the selected classes participated in the questionnaire survey, completing the instruments independently. The study yielded 3,275 valid questionnaires, distributed across educational stages as follows: 1,064 from grades 1–2, 1,069 from grades 3–4, and 1,142 from grades 5–6. Informed consent was obtained from all participants, and the study received ethical approval (Table 1).

Operationalization of the PL Model and Assessment

This study developed three grade-specific versions of the scale to align with children's cognitive development levels: PLS-Grade1–2 (physical literacy scale for elementary school students in grades 1–2), PLS-Grade 3–4 (physical literacy scale for elementary school students in grades 3–4), and PLS-Grade5–6 (physical literacy scale for elementary school students in grades 5–6). The scale's framework was constructed based on four fundamental dimensions of children's HL derived from existing Chinese policy documents and guidelines: knowledge, physical participation, physical competence, and physical motivation (Table 2).

Each version of the scale consists of two components: a Personal Information Questionnaire and a Physical Literacy Measurement Scale. The Personal Information Questionnaire comprises eight items collecting data on student demographics and behavioral characteristics: name, gender, age, grade, ethnicity, nearsightedness status, physical activity level, and internet usage patterns. The Physical Literacy Measurement Scale evaluates the four dimensions of children's PL using a 100-point scoring system. Dimensional weights were predetermined through Delphi expert consultation to ensure appropriate score

TABLE 1. Sociodemographic characteristics of the students (*N*=3,275).

Demographics	Frequency			Missing
	Grade 1–2 *n* (%)	Grade 3–4 *n* (%)	Grade 5–6 *n* (%)	
Gender				82
Boys	517 (49.6)	514 (49.2)	579 (52.3)	
Girls	525 (50.4)	530 (50.8)	528 (47.7)	
Grade				32
1/3/5	461 (43.9)	431 (40.8)	527 (46.4)	
2/4/6	590 (56.1)	625 (59.2)	609 (53.6)	
Ethnicity				41
Han	898 (86.1)	953 (89.7)	1,009 (89.4)	
Minority	145 (13.9)	109 (10.3)	120 (10.6)	
Nearsightedness				189
Yes	87 (8.2)	266 (26.5)	387 (38.1)	
No	897 (84.3)	667 (66.4)	572 (56.2)	
Unawareness	80 (7.5)	72 (7.1)	58 (5.7)	
Total	1,064 (100)	1,069 (100)	1,142 (100)	

TABLE 2. Summary structure and content classification of the three-vision scales.

Dimension	PLS-Grade	Items	Answer
Physical knowledge	1–2	B1–B10	Yes/No
	3–4	B1–B10	
	5–6	B1–B15	
Physical participation	1–2	D01C–D10C	Likert four-level scale (Never/1 to 3 times/4 to 6 times /Everyday)
	3–4	D01C–D12C	
	5–6	C01C–C14C	
		C01D–C14D	Likert five-level scale (Less than 0.5 hours/ 0.5 to 1 hour/1 to 2 hours/ more than 2 hours)
Physical motivation	1–2	C1–C10	Yes/No
	3–4	C1–C12	
	5–6	D1–D15	Likert five-level scale
Physical competence	1–2	D01B–D10B	Yes/No
	3–4	D01B–D12B	
	5–6	C01B–C14B	

Abbreviation: PLS=physical literacy scales.

allocation across components.

Physical Knowledge: The item banks were customized according to students' comprehension levels across different grades. For grades 1–4, 10 items were selected, while 15 items were chosen for grades 5–6. Each item consists of a true/false statement designed to assess children's understanding of physical activity, sedentary behavior recommendations, health perspectives, and safety awareness during physical activities (Table 3).

Physical Competence: The physical competence domain evaluates proficiency in various physical activities through capability-based questions (e.g., "Can you perform this activity?"). Following expert panel discussions, we developed grade-specific activity categories. For grades 1–2, we included 10 categories spanning leisure activities (e.g., shuttlecock kicking), moderate-intensity activities (cycling, gymnastics, roller skating, taekwondo), and high-intensity activities (dance, swimming, running, table tennis). The grades 3–4 questionnaire incorporated soccer and badminton as additional activities. For grades 5–6, we expanded the assessment to include sit-ups/pull-ups, mountain climbing, martial arts, and consolidated ball sports into two choice categories: one between basketball, soccer, or volleyball, and another between badminton, table tennis, or tennis. The final item counts were 10 for PLS-Grade1–2, 12 for PLS-Grade3–4, and 14 for PLS-Grade5–6.

Physical Participation

The physical participation domain evaluates weekly exercise activities across varying intensities and durations. Students in grades 1–4 respond using a four-point Likert scale, while grades 5–6 students answer questions structured on a four-point Likert scale. The number of items corresponds to the physical competences listed in the questionnaire: PLS-Grade 1–2 contains 10 items, PLS-Grade3–4 comprises 12 items, and PLS-Grade5–6 includes 14 items.

Physical Motivation

The Children's Self-Perception of Adequacy in and Predilection for Physical Activity (CSAPPA) Scale (*10*) was utilized to assess children's perceived competence in physical activities and their inclination toward

TABLE 3. Items for the physical knowledge and physical motivation dimensions.

Dimension	Grades	Item number	Item	Question type
Physical knowledge	1–2	B1–B2	The concept of physical activity	Yes/No
		B3	Benefits of physical activity	
		B4–B5	Knowledge about physical activity safety	
		B6	The concept of health	
		B7	Whether air pollution should continue to exercise	
		B8–B9	The dangers of sitting for a long time	
		B10	Knowledge about the amount of exercise	
	3–4	B1–B2	The concept of physical activity	Yes/No
		B3	Benefits of physical activity	
		B4–B5	Knowledge about physical activity safety	
		B6	The concept of health	
		B7	Whether to exercise under air pollution	
		B8–B9	The dangers of sitting for a long time	
		B10	Knowledge about the amount of exercise	
	5–6	B1–B2	The concept of physical activity	Yes/No
		B3	Benefits of physical activity	
		B4, B6–B7	Knowledge about physical activity safety	
		B5	Type of physical activity	
		B8	Knowledge about the amount of exercise	
		B9–B10	Whether to exercise under air pollution	
		B11–B14	The dangers of sitting for a long time	
		B15	The concept of health	

Continued

Dimension	Grades	Item number	Item	Question type
Physical motivation	1–2	C1	In order to strengthen physical fitness and get sick less	Judgment question
		C2	To bring me joy	
		C3	To learn new sport skills	
		C4	Like to meet new challenges	
		C5	To maintain good health	
		C6	In order to complete the sports test in school	
		C7	To meet new friends	
		C8	To look better for my own appearance	
		C9	Due to its inherent amusement	
		C10	Because exercise is important for me	
	3–4	C1	In order to strengthen physical fitness, get sick less	Judgment question
		C2	To bring me joy	
		C3	To learn new sport skills	
		C4	Like to meet new challenges	
		C5	To maintain good health	
		C6	In order to complete the sports test in school	
		C7	My friend wants me to exercise more	
		C8	To look better for my own appearance	
		C9	Due to its inherent amusement	
		C10	To meet new friends	
		C11	In order to get good results in sports tests of school	
		C12	Because exercise is important for me	
	5–6	D1	Because it is in line with life goals	Likert five-level scale
		D2	To relieve stress	
		D3	To control weight and improve body shape	
		D4	To bring me joy	
		D5	Because I enjoy of the process of exercise	
		D6	To maintain health	
		D7	To look better for my own appearance	
		D8	To maintain relationships	
		D9	To learn new sport skills	
		D10	Because everyone thinks I should exercise	
		D11	Because of the requirements of teachers and parents	
		D12	In order to get good results in sports tests	
		D13	Because of the deep love for sports	
		D14	Due to its inherent amusement	
		D15	Because of the importance of health	

participation. All scale items were adapted into age-appropriate language to ensure comprehension by adolescent participants. The PLS-Grade1–2 and PLS-Grade3–4 utilize true/false questions, with 10 questions per grade level. The PLS-Grade5–6 employs a five-point Likert scale comprising 15 questions (Table 3).

Statistical Analysis

To ensure data quality and independence, each class was assigned a dedicated investigator for the duration of the study. Students completed the questionnaires independently without teacher influence, while both the survey administrator and class teacher supervised the process to verify complete and accurate completion

of all questionnaires.

The measurement methodology was primarily guided by item response theory (IRT). Analysis was conducted using Winsteps software (version 3.66.0; https://winsteps.com/index.htm) to evaluate both dichotomous and multi-classification items. Items that failed to meet Rasch model criteria were either adjusted or eliminated as necessary.

RESULTS

Item Summary Statistics

Analysis of participant ability difficulty scores revealed mean values of −1.65, −1.06, −1.26, and −0.27, indicating that the items presented considerable challenge to participants. The Rasch model evaluation, utilizing Infit MNSQ and Outfit MNSQ average values, demonstrated optimal overall fit with values consistently falling between 0.5 and 1.5. These results indicate strong alignment between the three questionnaires and the ideal model, confirming robust data consistency (Table 4).

Unidimensionality

The variance explained by the measures across the three questionnaires was 53.1%, 50.3%, and 54.7%, respectively, all exceeding the 50% threshold. These results confirm unidimensionality, indicating that the items within each scale effectively measure a single, cohesive domain.

Item Person Map

The distribution of item difficulty and participant ability levels varied across grade groups. For grades 1–2, item difficulty spanned from −3 to 3 logit units, while participant ability levels ranged from -3 to 5 logit units. In grades 3–4, item difficulty ranged from -4 to 3 logit units, with participant ability levels distributed between −2 and 5 logit units. For grades 5–6, item difficulty extended from −4 to 3 logit units, while participant ability levels ranged from −2 to 4 logit units (Figure 1A–C). Across all three questionnaires, the item difficulty distribution demonstrated optimal alignment with participant ability levels.

Item Fit Statistical Analysis

Analysis of item fit statistics revealed robust measurement properties across all three scales. For PLS-Grade1–2, the Infit mean square statistics ranged from 0.82 to 1.31 with a mean of 1.00, while Outfit mean square values spanned from 0.62 to 1.78,

TABLE 4. Item summary statistics (*N*=3,275).

Psychometric attribute	PLS-Grade1–2	PLS-Grade3–4	PLS-Grade5–6
Number of item	40	46	72
Measure (Person)	1.68	1.14	0.40
Measure (Item)	0	0	0
Reliability			
Person reliability	0.81	0.81	0.89
Item reliability	1.00	1.00	1.00
Separation			
Person separation	2.04	2.09	2.86
Item separation	14.23	15.41	21.25
Item-fit statistics			
Infit MNSQ			
Mean square	1.00	1.01	1.11
Standard Deviation	0.10	0.09	0.46
ZSTD	−0.1	0.0	0.4
Outfit MNSQ			
Mmean square	1.04	1.12	1.17
Standard Deviation	0.28	0.29	0.56
ZSTD	0.3	0.8	0.5

Abbreviation: PLS=physical literacy scales; MNSQ=mean square; ZSTD=Z-standardized mean.

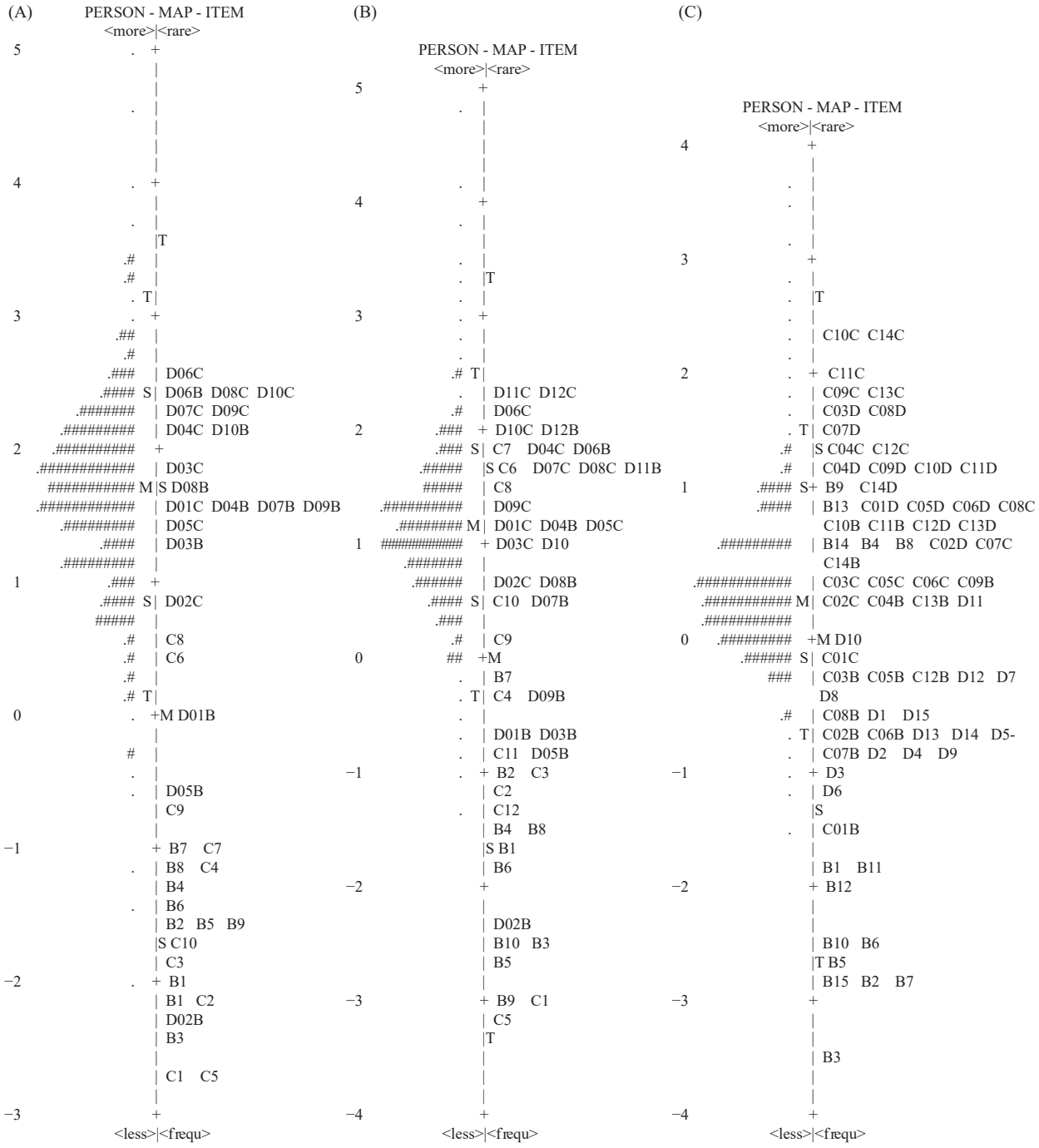

FIGURE 1. The distribution map of item difficulty and participant ability levels of (A) PLS-Grade1–2; (B) PLS-Grade3–4; and (C) PLS-Grade5–6.
Note: The left side of the chart shows the participants' ability levels and the right side displays the item difficulty levels. Each "#" means 8. Each "." means 1 to 7.
Abbreviation: PLS=physical literacy scales.

averaging 1.04. PLS-Grade3–4 demonstrated Infit MNSQ values between 0.79 and 1.34 (mean=1.01) and Outfit MNSQ values from 0.81 to 2.16 (mean=1.12). For PLS-Grade5–6, Infit MNSQ values ranged from 0.32 to 2.18 (mean=1.11), with corresponding Outfit MNSQ values spanning 0.37 to 2.54 (mean=1.17) (Supplementary Table S1, available at https://weekly.chinacdc.cn/).

DISCUSSION

The validation of questionnaire quality yielded

robust results. The item reliability coefficients across all four grade levels achieved a value of 1, with item separation values substantially exceeding 3. These findings demonstrate exceptional internal consistency within the questionnaire and validate the hierarchical structure of item difficulty levels. Furthermore, the questionnaire effectively discriminates among students with varying proficiency levels. The unidimensionality test revealed that all items across the three questionnaires met the necessary criteria, justifying subsequent Rasch model analysis. The item-person map demonstrates optimal alignment between the average item difficulty and respondents' ability levels, with items of varying difficulty distributed evenly across the scale. This distribution effectively accommodates primary school students across the spectrum of health literacy levels. Statistical analyses of all three questionnaires indicated favorable overall fit, showing strong concordance with the ideal model.

From an implementation perspective, our approach effectively addresses the varying cognitive capacities among primary school students while enabling precise assessment of their PL development. The questionnaires were strategically designed with grade-appropriate item counts, incorporating more questions for higher-grade students to maintain engagement throughout the measurement process. To accommodate younger students' developing comprehension abilities, their questionnaires featured simplified formats, such as true/false questions. Conversely, five-point Likert scale items were implemented for higher-grade students to comprehensively capture physical activity motivation factors. Additionally, the questionnaires employed grade-specific language to address variations in comprehension levels across age groups. This PL measurement scale, grounded in a thorough understanding of student characteristics, enhances both scientific rigor and practical utility through its hierarchical design and differentiated approach. These methodological insights offer valuable guidance for future PL measurement tool development.

Our study has several limitations. First, the reliance on primary school students' voluntary responses may introduce inherent participant subjectivity biases (*11*). Future studies should consider diverse data collection methods, such as parental or teacher assistance in questionnaire interpretation, to mitigate potential cognitive limitations-related subjectivity. Second, our use of cluster sampling may have resulted in population underrepresentation (*12*). We recommend that future research enhance survey methodology to reduce bias, moderately decrease scale difficulty, and expand sample size.

Conflicts of interests: No conflicts of interest.

Acknowledgments: We are grateful to all participants of primary school in Luzhou, Sichuan province; Cangzhou, Hebei province; Haidong, Hainan province; Shenzhen for their time and insights. We also want to express gratitude to the leaders and teachers in these schools and our postgraduate research assistants at Peking Union Medical College for their cooperation and assistance. Thanks to the staff of the regional Health Commission for their personnel and financial support.

Ethical statements: This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all procedures involving research study participants were reviewed and approved by the Institutional Review Board (IRB) of the Chinese Academy of Medical Sciences and Peking Union Medical College. Approval was granted with protocol number CAMS&PUMC-IEC-2022-026. Obtained written informed consent from all study participants. Obtained consent from parents or guardians.

Funding: This project is funded by the National Science Fund for Distinguished Young Scholars with the project number 71904205 and the Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences – A Strategic Study on Healthy China Development and Health system Reform (2021-I2M-1-046).

doi: 10.46234/ccdcw2025.049

[#] Corresponding author: Lili You, yll@pumc.edu.cn.

[1] School of Health Policy and Management, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China.

Submitted: September 20, 2024
Accepted: January 06, 2025
Issued: February 28, 2025

REFERENCES

1. Sørensen K, van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80. https://doi.org/10.1186/1471-2458-12-80.
2. Whitehead M. Definition of physical literacy and clarification of related issues. ICSSPE Bull 2013;65:29 – 35.
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 – 83. https://doi.org/10.1007/s40279-019-01063-3.
4. Tremblay MS, Costas-Bradstreet C, Barnes JD, Bartlett B, Dampier D, Lalonde C, et al. Canada's physical literacy consensus statement: process and outcome. BMC Public Health 2018;18(Suppl 2):1034. https://doi.org/10.1186/s12889-018-5903-x.
5. Yuan YB, Ji XY, Yang XM, Wang C, Samsudin S, Omar Dev RD. The effect of persistence of physical exercise on the positive psychological emotions of primary school students under the STEAM education concept. Int J Environ Res Public Health, 2022;19(18):11451. https://doi.org/10.3390/ijerph191811451.
6. Smith JJ, Eather N, Morgan PJ, Plotnikoff RC, Faigenbaum AD, Lubans DR. The health benefits of muscular fitness for children and adolescents: a systematic review and meta-analysis. Sports Med 2014;44(9):1209 – 23. https://doi.org/10.1007/s40279-014-0196-4.
7. Biddle SJH, Asare M. Physical activity and mental health in children and adolescents: a review of reviews. Br J Sports Med 2011;45(11):886 – 95. https://doi.org/10.1136/bjsports-2011-090185.
8. Song C, Ge S, Xue JJ, Yao WX. Physical exercise of primary and middle school students from the perspective of educational psychology and parents' entrepreneurship education. Front Psychol 2022;12:777069. https://doi.org/10.3389/fpsyg.2021.777069.
9. De Leeuw E. The political ecosystem of health literacies. Health Promot Int 2012;27(1):1 – 4. https://doi.org/10.1093/heapro/das001.
10. Cairney J, Veldhuizen S, Kurdyak P, Missiuna C, Faught BE, Hay J. Evaluating the CSAPPA subscales as potential screening instruments for developmental coordination disorder. Arch Dis Child 2007;92(11):987 – 91. https://doi.org/10.1136/adc.2006.115097.
11. Xue J, Liu YB, Sun KX, Wu LF, Liao K, Xia Y, et al. Validation of a newly adapted Chinese version of the newest vital sign instrument. PLoS One 2018;13(1):e0190721. https://doi.org/10.1371/journal.pone.0190721.
12. Fleary SA, Freund KM, Nigg CR. Development and validation of assessments of adolescent health literacy: a Rasch measurement model approach. BMC Public Health 2022;22(1):585. https://doi.org/10.1186/s12889-022-12924-4.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE S1. Outfit and Infit mean square statistics for individual items.

PLS-Grade	Item	Measure	SE	INFIT		OUTFIT		PT-Measure corr.
				MNSQ	ZSTD	MNSQ	ZSTD	
1–2	D06C	2.63	0.04	0.82	−3.4	0.73	−3.4	0.58
	D10C	2.48	0.04	1.28	5.2	1.62	6.7	0.44
	D08C	2.43	0.04	0.92	−1.6	1.01	0.1	0.52
	D06B	2.43	0.07	0.91	−3.4	0.92	−2.0	0.41
	D07C	2.29	0.03	0.88	−2.7	0.93	−0.9	0.56
	D09C	2.27	0.03	0.86	−3.4	0.94	−0.8	0.58
	D10B	2.14	0.07	1.00	0.2	1.04	1.2	0.31
	D04C	2.11	0.03	1.29	6.8	1.31	4.4	0.46
	D03C	1.92	0.03	1.31	7.7	1.47	6.9	0.46
	D08B	1.70	0.07	0.98	−0.8	0.98	−0.6	0.34
	D07B	1.64	0.07	0.94	−3.4	0.92	−3.0	0.40
	D04B	1.61	0.07	0.99	−0.4	1.0	−0.2	0.33
	D01C	1.57	0.03	1.09	2.3	1.17	2.8	0.51
	D09B	1.54	0.07	0.90	−5.0	0.88	−4.5	0.44
	D05C	1.36	0.03	0.92	−2.1	0.97	−0.5	0.54
	D03B	1.24	0.07	0.94	−2.9	0.93	−2.3	0.40
	D02C	0.86	0.04	0.98	−0.4	1.26	2.9	0.52
	C8	0.50	0.08	1.01	0.4	1.02	0.4	0.28
	C6	0.38	0.08	1.11	2.7	1.41	6.4	0.12
	D01B	−0.04	0.09	1.03	0.5	1.00	0	0.25
	D05B	−0.52	0.10	0.97	−0.4	0.90	−1.0	0.29
	C9	−0.78	0.11	0.97	−0.3	0.88	−1.0	0.28
	C7	−0.97	0.12	0.98	−0.2	0.89	−0.9	0.26
	B7	−1.07	0.12	1.04	0.5	1.49	3.1	0.13
	B8	−1.13	0.13	0.98	−0.1	0.94	−0.4	0.23
	C4	−1.20	0.13	0.95	−0.4	0.84	−1.1	0.28
	B4	−1.35	0.14	0.99	0	1.10	0.6	0.2
	B6	−1.49	0.14	1.07	0.6	1.78	3.8	0.02
	B2	−1.56	0.15	0.95	−0.4	0.68	−2.0	0.28
	B5	−1.56	0.15	1.05	0.4	1.40	2.1	0.10
	B9	−1.63	0.15	0.98	−0.1	0.97	−0.1	0.23
	C10	−1.65	0.15	0.92	−0.6	0.69	−1.8	0.30
	C3	−1.83	0.17	0.94	−0.3	0.63	−2.1	0.27
	B1	−1.98	0.18	1.02	0.2	1.09	0.5	0.13
	C2	−2.08	0.19	0.95	−0.2	0.90	−0.4	0.21
	B10	−2.15	0.19	0.98	0	1.69	2.5	0.15
	D02B	−2.31	0.21	0.97	−0.1	0.90	−0.3	0.17
	B3	−2.39	0.21	1.02	0.2	0.91	−0.2	0.13
	C1	−2.71	0.25	0.94	−0.2	0.62	−1.3	0.21
	C5	−2.71	0.25	0.94	−0.2	0.79	−0.6	0.19

Continued

PLS-Grade	Item	Measure	SE	INFIT		OUTFIT		PT-Measure corr.
				MNSQ	ZSTD	MNSQ	ZSTD	
3–4	D12C	2.4	0.05	1.34	5.2	1.02	2.7	0.49
	D11C	2.27	0.04	1.06	1.0	0.96	−0.5	0.55
	D06C	2.25	0.04	0.96	−0.8	0.87	−1.9	0.59
	D12B	2.07	0.07	0.98	−0.7	0.97	−0.8	0.33
	D10C	2.05	0.04	0.87	−2.6	0.93	−1.1	0.59
	D04C	1.89	0.04	1.20	4.0	1.18	3.0	0.47
	D06B	1.87	0.07	0.95	−2	0.91	−2.5	0.38
	C7	1.77	0.07	1.04	1.5	1.05	1.6	0.24
	D07C	1.69	0.04	0.93	−1.8	0.93	−1.3	0.54
	C6	1.68	0.07	1.07	2.7	1.11	3.6	0.20
	D08C	1.66	0.04	0.86	−3.4	0.84	−3.2	0.60
	D11B	1.61	0.07	0.97	−1.3	0.97	−1.2	0.34
	C8	1.55	0.07	1.10	4.7	1.13	4.6	0.15
	D09C	1.29	0.03	0.79	−6.1	0.81	−4.4	0.58
	D05C	1.21	0.03	0.84	−4.7	0.85	−3.4	0.54
	D04B	1.17	0.07	1.04	2.2	1.04	1.6	0.25
	D01C	1.10	0.03	0.98	−0.4	1.09	2.0	0.48
	D10B	1.00	0.07	0.95	−3.1	0.94	−2.5	0.37
	D03C	0.96	0.03	1.19	5.1	1.26	5.3	0.45
	D08B	0.67	0.07	0.96	−1.7	0.96	−1.2	0.34
	D02C	0.61	0.03	0.97	−0.8	1.01	0.2	0.44
	D07B	0.51	0.07	1.02	0.8	1.04	1.2	0.25
	C10	0.43	0.07	1.00	0.1	1.05	1.3	0.27
	C9	0.13	0.07	1.02	0.6	1.09	1.8	0.23
	B7	−0.12	0.08	1.11	2.7	1.27	4.6	0.07
	C4	−0.40	0.08	0.99	−0.2	1.08	1.2	0.24
	D09B	−0.41	0.08	0.95	−1.1	0.95	−0.8	0.31
	D01B	−0.62	0.09	1.03	0.5	1.15	1.9	0.18
	D03B	−0.66	0.09	0.97	−0.5	0.98	−0.2	0.26
	C11	−0.79	0.09	1.01	0.1	1.04	0.5	0.20
	D05B	−0.79	0.09	0.98	−0.2	1.00	0	0.23
	C3	−0.94	0.1	0.99	−0.1	1.03	0.4	0.21
	B2	−1.00	0.1	1.04	0.6	1.15	1.5	0.13
	C2	−1.09	0.1	0.97	−0.4	0.96	−0.3	0.24
	C12	−1.27	0.11	0.98	−0.2	0.95	−0.4	0.21
	B8	−1.44	0.12	1.05	0.5	1.74	5.0	0.04
	B4	−1.52	0.12	1.07	0.7	1.85	5.3	−0.02
	B1	−1.67	0.13	1.00	0.1	1.24	1.6	0.13
	B6	−1.81	0.13	1.02	0.2	1.05	0.4	0.12
	D02B	−2.41	0.17	1.00	0	1.27	1.3	0.09
	B3	−2.47	0.18	1.02	0.2	1.19	0.9	0.08
	B10	−2.50	0.18	0.98	0	0.9	−0.4	0.16

Continued

PLS-Grade	Item	Measure	SE	INFIT		OUTFIT		PT-Measure corr.
				MNSQ	ZSTD	MNSQ	ZSTD	
	B5	−2.72	0.2	1.03	0.2	2.07	3.6	−0.01
	B9	−2.99	0.23	1.02	0.2	2.16	3.4	−0.03
	C1	−2.99	0.23	0.99	0	0.82	−0.6	0.13
	C5	−3.22	0.25	1.0	0.1	1.34	1.2	0.07
5–6	C14C	2.31	0.06	1.15	2	1.04	0.6	0.48
	C10C	2.28	0.06	1.27	3.5	1.13	1.7	0.50
	C11C	2.04	0.05	1.32	4.4	1.33	4.2	0.46
	C09C	1.88	0.05	1.28	4.0	1.24	3.3	0.46
	C13C	1.86	0.05	0.99	−0.1	0.97	−0.5	0.53
	C08D	1.66	0.05	1.18	3.4	1.17	3.2	0.34
	C03D	1.64	0.05	1.20	3.7	1.18	3.5	0.35
	C07D	1.44	0.05	1.19	3.4	1.17	3.2	0.41
	C04C	1.35	0.04	1.38	6.5	1.40	6.3	0.29
	C12C	1.34	0.04	0.77	−4.8	0.80	−3.9	0.50
	C11D	1.22	0.05	0.65	−7.4	0.64	−7.7	0.34
	C09D	1.19	0.05	0.72	−5.7	0.71	−6.2	0.34
	C10D	1.13	0.05	0.69	−6.6	0.66	−7.1	0.34
	C04D	1.13	0.05	1.13	2.4	1.12	2.2	0.22
	C14D	1.08	0.04	0.64	−7.8	0.61	−8.4	0.35
	B9	1.02	0.05	0.58	−9.9	0.61	−9.9	0.02
	C13D	0.91	0.04	0.75	−5.3	0.73	−5.6	0.40
	C01D	0.9	0.04	1.19	3.5	1.18	3.2	0.48
	C06D	0.89	0.04	1.16	3	1.17	3.1	0.42
	C10B	0.88	0.04	1	−0.1	0.99	−0.2	0.39
	C11B	0.84	0.04	0.98	−0.5	0.97	−0.5	0.4.0
	B13	0.84	0.05	0.42	−9.9	0.45	−9.9	0.13
	C08C	0.81	0.03	0.8	−5.5	0.79	−5.0	0.49
	C05D	0.77	0.04	1.17	3.2	1.14	2.7	0.44
	C12D	0.77	0.04	1.00	0	1.00	0	0.41
	C02D	0.67	0.04	1.12	2.5	1.12	2.2	0.39
	B4	0.67	0.05	0.34	−9.9	0.39	−9.9	0.01
	B14	0.66	0.05	0.32	−9.9	0.37	−9.9	0.05
	C14	0.66	0.03	1.05	1.8	1.08	1.6	0.34
	B8	0.65	0.05	0.34	−9.9	0.4	−9.9	−0.05
	C07C	0.62	0.03	0.80	−5.8	0.8	−5.3	0.41
	C09B	0.55	0.03	1.07	2.8	1.11	2.1	0.32
	C06C	0.54	0.03	0.80	−6.1	0.8	−5.3	0.47
	C05C	0.53	0.03	0.93	−2.1	0.92	−2.0	0.48
	C03C	0.48	0.03	1.33	8.6	1.35	8.1	0.26
	C02C	0.39	0.03	1.05	1.6	1.07	1.8	0.37
	C04B	0.35	0.03	1.17	7.5	1.41	6.2	0.19
	C13B	0.35	0.03	1.04	1.9	1.06	1.0	0.33

Continued

PLS-Grade	Item	Measure	SE	INFIT		OUTFIT		PT-Measure corr.
				MNSQ	ZSTD	MNSQ	ZSTD	
	D11	0.29	0.03	1.21	5.5	1.28	6.8	0.35
	D10	−0.03	0.03	1.04	0.9	1.09	2.1	0.39
	C01C	−0.12	0.03	0.79	−6.9	0.83	−4.5	0.39
	D7	−0.25	0.03	1.14	3.2	1.28	5.6	0.35
	D12	−0.25	0.03	1.12	2.8	1.21	4.3	0.37
	C03B	−0.31	0.04	1.12	3.0	1.93	7.6	0.14
	C05B	−0.34	0.04	1.00	−0.1	0.89	−1.1	0.31
	D8	−0.35	0.03	1.06	1.5	1.08	1.7	0.41
	C12B	−0.35	0.04	1.00	0.1	1.01	0.2	0.29
	D1	−0.54	0.03	0.74	−6.1	0.73	−5.8	0.51
	C08B	−0.57	0.04	0.98	−0.3	1.28	2.2	0.27
	D15	−0.58	0.03	0.86	−3.0	0.89	−2.2	0.46
	D5	−0.66	0.03	0.85	−3.3	0.82	−3.6	0.49
	C02B	−0.68	0.05	1.05	0.8	1.21	1.6	0.19
	D14	−0.69	0.04	0.86	−3.0	0.86	−2.6	0.50
	D13	−0.70	0.04	0.84	−3.3	0.83	−3.2	0.51
	C06B	−0.73	0.05	0.99	−0.2	0.88	−0.9	0.26
	D9	−0.75	0.04	0.85	−3.1	0.89	−2.1	0.43
	D4	−0.76	0.04	0.78	−4.6	0.73	−5.4	0.51
	D2	−0.80	0.04	0.67	−7.1	0.65	−7.0	0.51
	C07B	−0.92	0.06	1.02	0.3	1.1	0.7	0.18
	D3	−1.08	0.04	0.86	−2.5	0.82	−3.0	0.43
	D6	−1.11	0.04	0.79	−3.9	0.74	−4.5	0.43
	C01B	−1.55	0.09	0.97	−0.1	1.35	1.3	0.14
	B11	−1.85	0.07	2.18	9.9	2.5	9.9	0.08
	B1	−1.87	0.07	2.18	9.9	2.34	9.9	0.10
	B12	−2.06	0.07	2.08	9.9	2.2	9.9	0.18
	B6	−2.48	0.08	2.13	9.9	2.54	9.9	0.05
	B10	−2.50	0.08	2.09	9.9	2.30	9.9	0.10
	B5	−2.65	0.09	2.08	9.9	2.35	9.9	0.09
	B7	−2.81	0.09	2.07	9.9	2.35	9.9	0.09
	B2	−2.88	0.10	2.09	9.9	2.53	9.9	0.04
	B15	−2.88	0.10	2.07	9.9	2.44	9.9	0.07
	B3	−3.48	0.12	2.01	7.0	2.33	7.4	0.11

Abbreviation: PLS=physical literacy scales; SE=standard error; PT-Measure corr.=point measure correlation; MNSQ=mean square; ZSTD=Z-standardized mean.