informatics

MDPI

Article

Development and Usability Study of Mobile Physical Fitness Assessment for Lower Secondary Students Aged 13–15 Years

Nirut Futhep [1], Jirakrit Leelarungrayub [1,2] and Teerawat Kamnardsiri [1,3,*]

[1] Department of Sports Science, Multidisciplinary and Interdisciplinary School, Chiang Mai University, Chiang Mai 50200, Thailand; nirut_f@cmu.ac.th (N.F.); donrawee.leela@cmu.ac.th (J.L.)
[2] Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand
[3] Department of Digital Game, College of Arts, Media, and Technology, Chiang Mai University, Chiang Mai 50200, Thailand
* Correspondence: teerawat.k@cmu.ac.th; Tel.: +66-81-616-5860

Abstract

Physical fitness assessment is an essential component of school-based physical education, yet in many schools it still relies on manual recording that is time-consuming and prone to error. This study aimed to (1) develop a web-based application, accessible on computers and mobile devices, to assist physical education teachers in assessing the physical fitness of lower secondary school students aged 13–15 years, and (2) evaluate user satisfaction and the usability of the application. A Research and Development (R&D) approach based on the System Development Life Cycle (SDLC) was employed. The study included 36 physical education teachers from secondary schools in Mueang District, Chiang Mai Province, Thailand, who used the application to record and analyze five physical fitness components—Body Mass Index (BMI), Sit-and-Reach, 30-s Modified Push-ups, 60-s Sit-ups, and the 3-min Step Test—using 1580 student records. Instruments included the developed application, a user satisfaction questionnaire, and the System Usability Scale (SUS); data were analyzed using descriptive statistics. A comparative analysis of the sit-and-reach test showed a high level of consistency between manual assessments and the mobile application, indicating that the system classified students' flexibility performance consistently with the manual criteria. The mean SUS score was 87.08, exceeding the acceptable usability threshold, although the internal consistency of the Thai SUS in this sample was low (Cronbach's $\alpha = 0.42$); participants reported high to very high satisfaction across all evaluation dimensions. The application's automatic classifications were identical to the manual classifications across all fitness tests analyzed, indicating that it implements the scoring criteria correctly. Overall, the application reproduced the manual classifications and teachers rated its usability as excellent, suggesting that it is a promising and practical tool for physical education teachers in school settings; its effects on recording time and data-entry errors were not quantified in this study.

Keywords: health promotion; educational technology; mobile application; physical fitness assessment; system usability scale; information technology

Check for updates

Received: 23 September 2026
Revised: 29 October 2026
Accepted: 2 November 2026
Published: 3 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Physical fitness is a fundamental component of health and well-being, particularly during adolescence, a critical period of physical, psychological, and social development. It is defined as the ability to perform daily activities efficiently while maintaining sufficient

energy for leisure activities and unexpected demands [1–5]. Physical fitness comprises health-related components, including cardiorespiratory endurance, muscular strength, muscular endurance, flexibility, and body composition, as well as skill-related components such as agility, balance, coordination, power, speed, and reaction time [6–8]. A high level of physical fitness during adolescence is associated with numerous health benefits, including reduced risks of non-communicable diseases, improved mental health, enhanced cognitive function, and better academic achievement [9–13]. Because physical activity habits established during adolescence often persist into adulthood, promoting physical fitness at this stage is considered an important public health priority. Schools provide an ideal setting for promoting physical activity because they offer structured opportunities for students to develop healthy behaviors through physical education (PE). Evidence indicates that school-based PE interventions, including comprehensive physical activity programs and high-intensity interval training (HIIT), significantly improve students' cardiorespiratory fitness, muscular strength, physical activity participation, psychological well-being, and academic performance [14–18]. Furthermore, supportive school environments and positive social influences from teachers and peers enhance students' engagement in physical activity [19,20]. Consistent with these findings, Thailand's National Education Act and the Basic Education Core Curriculum emphasize the promotion of students' physical fitness through Health and Physical Education. The curriculum requires lower secondary school students to assess, monitor, and improve their physical fitness according to national standards [21]. At The Prince Royal's College, Chiang Mai, Thailand, lower secondary students undergo annual physical fitness assessments based on the standards established by the Department of Physical Education. However, the current assessment process relies primarily on manual recording and interpretation of test results. Given the large number of students, this approach is time-consuming, susceptible to human error, and inefficient for data management and long-term monitoring. Previous studies have suggested that digital systems and mobile technologies can improve the efficiency, accuracy, and accessibility of physical fitness assessment by facilitating data recording, analysis, storage, and retrieval [22–26]. Mobile apps have already been used in school settings; for example, mandatory after-school use of step-tracker apps improved physical activity, body composition, and fitness in adolescents [27]. We therefore adopted a web-based architecture that runs in a standard browser on computers and mobile devices, so that teachers can use the system without installing software. Therefore, this study aimed to develop and evaluate a web-based, mobile-accessible physical fitness application for lower secondary school students. The application was designed to support the standardized recording and classification of fitness test results and to establish a centralized database for continuous monitoring and health promotion among students.

The study was guided by three research questions. RQ1: To what extent do the physical fitness classifications generated automatically by the developed application agree with the classifications produced manually by the teachers who administered the tests from the same measurements, using the national age- and sex-specific criteria? RQ2: How do physical education teachers rate the usability of the application according to the System Usability Scale (SUS), and how internally consistent is the scale in this sample? RQ3: How satisfied are physical education teachers with the application's functions for recording, summarizing, and managing students' physical fitness data? Section 2 reviews related work and identifies the research gap, Section 3 describes the design, materials, and procedures, and Sections 4–6 present the results, discussion, and conclusions.

2. Related Work

2.1. Digital Technologies in Physical Education

Digital technologies are increasingly integrated into physical education (PE), ranging from information and communication technologies to artificial intelligence. Recent systematic reviews describe their adoption in PE while cautioning that realizing pedagogical benefits depends on ethical frameworks and investment in teachers' professional development [28]; they also note that most earlier reviews focused on isolated technologies or specific instructional contexts [29]. Within this body of work, mobile applications for young people have been studied mainly as intervention tools. For example, the AIMFIT trial evaluated apps intended to improve fitness and physical activity among young people [24], and mandatory after-school use of step-tracker apps was associated with improvements in physical activity, body composition, and fitness in adolescents [27]. These studies concern behavior change rather than the standardized, teacher-administered fitness testing that the national curriculum requires [21].

2.2. Smartphone Tools for Assessing Fitness and Activity in Young People

Fewer studies address the use of smartphones to measure or assess the fitness and activity of young people. A systematic review of smartphone approaches to measuring physical activity in healthy young people identified only five eligible studies and reported that studies typically examined a single measurement property, with few addressing reliability or validity [30]. Test-specific apps have been evaluated more rigorously. The My Jump 2 app was compared with a photoelectric reference system (OptoJump) in 48 primary-school children aged 11–14 years and was reported to be a valid, reliable, and useful tool for measuring vertical jump [31]; however, it addresses a single test and does not manage records for a class or school. For multi-test batteries, Sulla-Torres et al. [32] developed a multiplatform computer-vision system for four school fitness tests (sit-and-reach, horizontal jump, biceps curl, and sit-ups) in 228 schoolchildren aged 8–18 years and reported 94% precision in field evaluations. This approach relies on video capture of each test and does not include the BMI, push-up, and step-test components used in the present battery. Standardized field batteries such as ALPHA [26] provide established protocols on which digital tools can build, and in Thailand the Department of Physical Education prescribes age- and sex-specific criteria for classifying performance [2] and provides a program for storing and reporting test results [4].

2.3. Research Gap and Contribution

Table 1 summarizes the studies discussed above. Three limitations are evident. First, most digital work with young people targets behavior change [24,27] or a single test [31], and the only multi-test system among those reviewed depends on video capture and covers a different set of tests [32]. Second, teacher-facing functions—recording the five-component battery, automatic classification against age- and sex-specific national criteria, and cohort-level data management—are not the focus of the studies reviewed. Third, the evidence on the measurement properties of smartphone approaches in young people remains limited [30].

The present study addresses these gaps by developing and evaluating a teacher-facing application that records the five tests of the national battery, classifies results automatically against the national criteria, and stores records centrally, and by examining its agreement with manual classification, its usability, and teacher satisfaction (RQ1–RQ3).

Table 1. Summary of selected related studies on digital and mobile tools for fitness and physical activity in young people.

Study	Context and Participants	Tool and Measures	Evaluation Reported	Scope Relative to the Present Study
Direito et al. [24]	Young people; randomized controlled trial	Smartphone apps to improve fitness and physical activity	Pragmatic randomized controlled trial (AIMFIT)	Behavior-change intervention; not a teacher-facing system for recording and classifying fitness tests
Mateo-Orcajada et al. [27]	Adolescents; mandatory after-school use	Step-tracker apps	Physical activity, body composition, and fitness outcomes	Behavior-change intervention; no standardized test battery or scoring criteria
Parmenter et al. [30]	Systematic review; healthy young people aged 10–24 years (5 eligible studies)	Smartphone approaches to measure physical activity	Review of reliability, validity, and responsiveness	Few studies; typically only one measurement property reported
Bogataj et al. [31]	48 primary-school children aged 11–14 years	My Jump 2 app; vertical jump (SJ, CMJ, CMJAM)	Concurrent validity against OptoJump; test–retest reliability (ICC); Bland–Altman analysis	Single test; no multi-test battery, classification, or record management
Sulla-Torres et al. [32]	228 schoolchildren aged 8–18 years	Multiplatform computer-vision system (smartphone video); sit-and-reach, horizontal jump, biceps curl, sit-ups	Precision against traditional evaluation (94% field, >95% laboratory); software quality model	Video-based; BMI, push-ups, and step test not included
This study	36 teachers; 1580 student records (13–15 years)	Teacher-facing application; BMI, sit-and-reach, push-ups, sit-ups, step test	Agreement with manual classification; SUS; teacher satisfaction	Recording, automatic classification against national criteria, and central data management

3. Materials and Methods

3.1. Research Design

This study employed a Research and Development (R&D) approach with the objective of developing a mobile application for physical education teachers to assess the physical fitness of lower secondary school students aged 13–15 years. In addition, the study examined user satisfaction and system usability of the developed application. The research was conducted following the System Development Life Cycle (SDLC) framework, which includes the stages of planning, design, development, testing, and system refinement. The application was developed as a web-based application that runs in a standard web browser on personal computers and on mobile devices such as tablets and iPads, so that no installation is required and teachers can use whichever device is available at the test site. Because it is used on mobile devices as well as computers, it is referred to as a mobile-accessible application in this paper. The scope of data used in this study included students' physical fitness test results based on standardized criteria established by the Department of Physical Education, Thailand. These comprised body mass index (BMI), flexibility, muscular strength and endurance, and cardiorespiratory endurance. In addition, data on user satisfaction and feedback were collected from physical education teachers who participated in the system trial.

Requirements analysis: Requirements were elicited through semi-structured interviews in which approximately 200 lower-secondary teachers of the Health and Physical Education learning area, from approximately 36 schools under the Ministry of Education in Chiang Mai Province, were asked about their needs for a system to assess students' physical fitness.

Development platform and tools: The system runs on personal computers and on mobile devices such as tablets and iPads through a standard web browser. PHP was used for the pages, authentication, API, and database access; HTML5 for page structure;

CSS/SCSS, including Bootstrap and the SB Admin template, for styling; JavaScript and jQuery for browser interactions, DOM handling, and AJAX requests to the API; and SQL with a MariaDB/MySQL database for data storage.

Data structure: Each test record in the application stores the student's first name and surname, sex, and age (in years), the measured value (for BMI, also height and weight), the resulting classification level, and a timestamp. Records are organized by test type (BMI, sit-and-reach, push-ups, sit-ups, and step test), so that they can be searched, edited, imported, and exported for later analysis (Appendix A, Figures A1–A7).

3.2. Participants

The sample group in this study consisted of 36 physical education teachers (75% male, aged 20–40 years) from secondary schools located in Mueang District, Chiang Mai Province. All participants were responsible for teaching at the lower secondary level. This study was conducted in accordance with ethical standards for human research and received approval from the Research Ethics Committee of the Multidisciplinary and Interdisciplinary College, Chiang Mai University (Approval No. MIdSEC-68EX-003, dated 12 January 2026).

In addition to the 36 physical education teachers, the study used physical fitness data of lower secondary school students aged 13–15 years. Each teacher tested the application by entering a data set of 40–46 students, giving 1580 records in total. Some teachers used the same student data set, so the 1580 records correspond to 1355 distinct students. The unit of analysis in Section 4.1 is the record entered by a teacher.

In this study, the sample size was determined using the formula proposed by Taro Yamane, as cited in Joskow [33]. The calculation was conducted at a 95% confidence level with a margin of error of ±5%. The target population (N) comprised 40 physical education teachers responsible for teaching lower secondary school students in Mueang District, Chiang Mai Province. Substituting N = 40 and e = 0.05 into Equation (1) as follows:

$$n = \frac{N}{1 + Ne^2} \quad (1)$$

where n = required sample size, N = population size, and e = margin of error (0.05). Based on Equation (1), the calculated sample size was n = 36.36, which was rounded up to a minimum required sample of 36 participants. This matches the number of physical education teachers (n = 36) recruited for this study. This formula provides a systematic approach to determining an appropriate sample size for studies requiring statistical representativeness under specified confidence and precision levels.

3.3. Materials and Instruments

3.3.1. The Mobile Application

The system architecture, designed using a layered architecture approach, is illustrated in Figure 1. The workflow of the system can be summarized as follows.

(1) Client Layer: The Client Layer serves as the user interaction layer, supporting access through web browsers, mobile applications, and third-party systems. Its primary function is to send client requests to the system and receive responses through the Response Handler. In cases of errors or failures, the Error Handler returns appropriate error messages to users.
(2) REST API Gateway: The REST API Gateway acts as the central entry point of the system. It receives incoming requests from the Client Layer, validates them, and routes them to the appropriate services within the Business Layer.
(3) Business Layer: The Business Layer contains the core business logic and system functionalities. It is divided into several service components, including Core Services such

as the BMI Service for Body Mass Index calculation, Section Service, and Login Service; Support Services such as the Menu Service and Report Service for menu management and report generation; and Auth Middleware, which is responsible for authentication and authorization before allowing access to sensitive data or services.

(4) Data Layer: The Data Layer manages data storage and communication with supporting systems. It consists of PHP Data Objects (PDO) Connection for connecting to the primary database system, such as MySQL; Cache, which stores frequently accessed data to improve system performance and reduce processing time; and Log Handler, which records system activities and events for monitoring and debugging purposes.

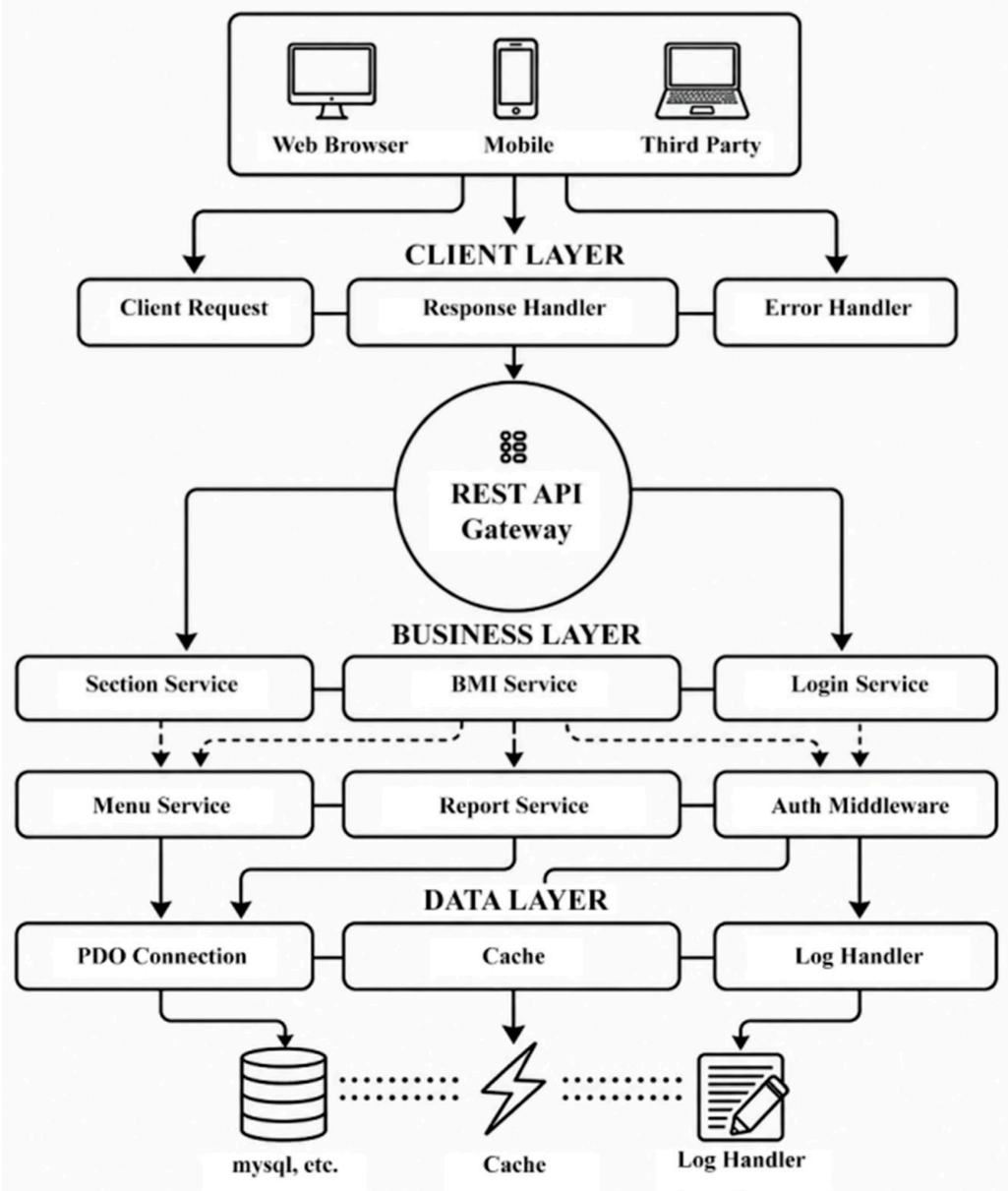
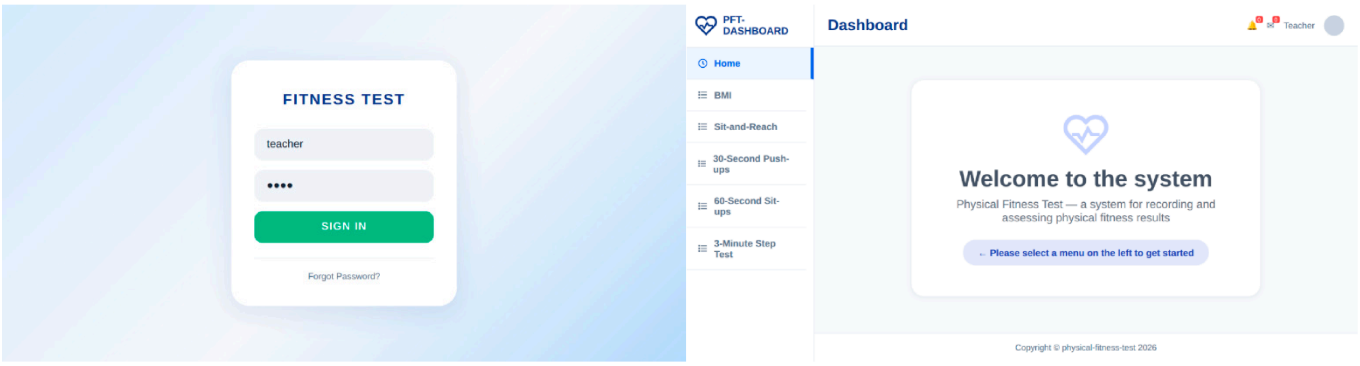


Figure 1. System architecture of the physical fitness test mobile application.

The user interface is shown in Figures 2–4, which are English-translated renderings of the Thai interface; further screens are provided in Appendix A (Figures A1–A7).

Access and navigation: Teachers sign in with a username and password (Figure 2a) and are taken to the dashboard (Figure 2b). The left-hand menu provides one module for each component of the battery—BMI, sit-and-reach, 30-s push-ups, 60-s sit-ups, and

the 3-min step test—and teachers move between modules by selecting the corresponding menu item.



(**a**) (**b**)

Figure 2. Access to the application: (**a**) login screen; (**b**) dashboard home screen with the module menu (English-translated rendering of the Thai interface).

Test recording and result calculation: Within a test module, the record table lists the entries and offers search, sorting, and paging controls together with three function buttons: Create New Entry, Import List, and Export List. To record a test, the teacher selects Create New Entry, chooses the student's sex, and enters the age in years, the first name and surname, and the measured value (Figure 3a). After the entry is saved, the application automatically classifies the result into one of five levels (very low to very good) using the age- and sex-specific criteria in Tables 2–4 and stores the result with a timestamp (Figure 3b); for example, a 13-year-old boy with a measured value of 156 in the 3-min step test is classified as "Good". The same procedure applies to the other test modules; for example, Figure 3c,d show the Create New Entry dialog and the saved, automatically classified record for the 60-s sit-ups.

Table 2. Standardized criteria for physical fitness assessment in 13-year-old male and female students.

Test Item	Very Good		Good		Moderate		Low		Very Low	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Sit-and-Reach	≥21	≥23	16–20	17–22	11–15	12–16	6–10	6–11	≤5	≤5
30-s Push-ups	≥34	≥30	27–33	24–29	20–26	18–23	12–19	11–17	≤12	≤10
60-s Sit-ups	≥47	≥41	38–46	33–40	28–37	24–32	19–27	16–23	≤18	≤15
3-min Step Test	≥169	≥152	150–168	140–151	124–149	118–139	99–123	97–117	≤98	≤96

Table 3. Standardized criteria for physical fitness assessment in 14-year-old male and female students.

Test Item	Very Good		Good		Moderate		Low		Very Low	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Sit-and-Reach	≥23	≥24	16–22	19–23	11–15	14–18	6–10	8–13	≤5	≤7
30-s Push-ups	≥36	≥31	28–35	25–30	21–27	18–24	14–20	12–17	≤13	≤11
60-s Sit-ups	≥49	≥43	39–48	34–42	30–38	25–33	20–29	16–24	≤19	≤15
3-min Step Test	≥171	≥160	155–170	147–159	130–154	124–146	105–129	101–123	≤104	≤100

Table 4. Standardized criteria for physical fitness assessment in 15-year-old male and female students.

Test Item	Very Good		Good		Moderate		Low		Very Low	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Sit-and-Reach	≥25	≥27	20–24	19–26	14–19	14–18	8–13	8–13	≤7	≤12
30-s Push-ups	≥38	≥34	30–37	27–33	23–29	20–26	15–22	13–19	≤14	≤12
60-s Sit-ups	≥50	≥45	40–49	36–44	31–39	27–35	21–30	19–26	≤20	≤18
3-min Step Test	≥173	≥165	156–172	148–164	131–155	125–147	105–130	102–124	≤104	≤101

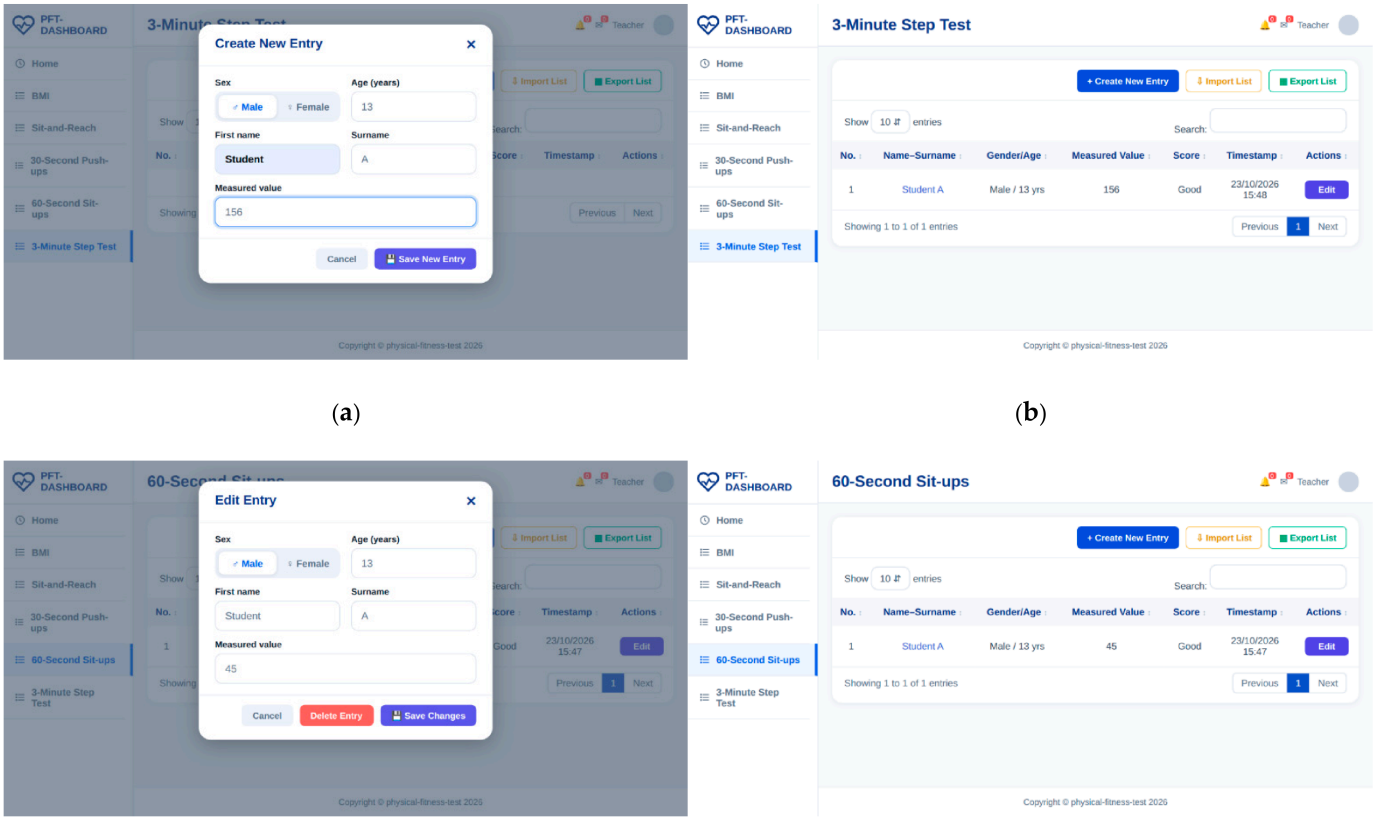
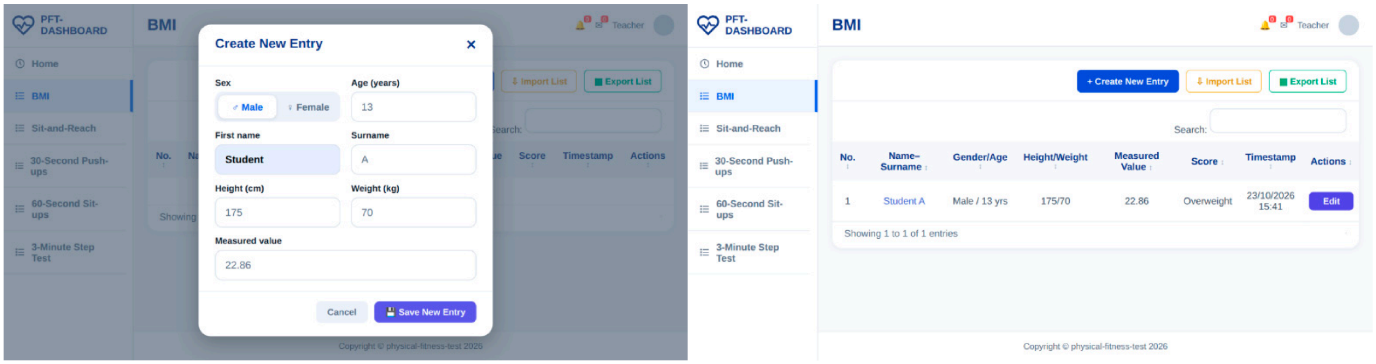


Figure 3. Recording and managing a test result: (**a**) Create New Entry dialog (3-min step test); (**b**) saved record with the automatically assigned classification of the 3-min step test; (**c**) Create New Entry dialog (60-s Sit-ups) and (**d**) saved record with the automatically assigned classification of the 60-s Sit-ups. Student names are fictitious (English-translated rendering of the Thai interface).

Data management: Selecting Edit opens a dialog in which the teacher can correct the sex, age, name, or measured value, delete the entry, or cancel (the Edit button is shown in Appendix A, Figure A3); the classification is then updated by the same criteria. Import List and Export List allow records to be imported and exported, and exported data are provided in Microsoft Excel format for further analysis (Section 3.4).

BMI module: In the BMI module the teacher enters the student's height (cm) and weight (kg) in addition to sex, age, and name; the application calculates BMI from them (for example, 175 cm and 70 kg give 22.86 kg/m^2) and assigns a weight-status category ("Overweight" in this example, because 22.86 lies between 21.61 and 25.76 for 13-year-old boys; Table 5), and the record table additionally displays the height and weight (Figure 4).



(**a**) (**b**)

Figure 4. BMI module: (**a**) Create New Entry dialog with height and weight; (**b**) saved record with the calculated BMI and weight-status category. Student names are fictitious (English-translated rendering of the Thai interface).

Table 5. Body mass index cut-offs (kg/m^2) for male and female students aged 13–15 years (source: Department of Physical Education [2]).

Age (Years)	Sex	Very Thin	Thin	Proportionate	Overweight	Obese
13	Male	≤12.02	12.03–17.42	17.43–21.60	21.61–25.76	≥25.77
13	Female	≤12.74	12.75–17.36	17.37–21.64	21.65–25.85	≥25.86
14	Male	≤12.53	12.54–17.65	17.66–21.95	21.96–26.26	≥26.27
14	Female	≤13.19	13.20–18.05	18.06–22.93	22.94–26.91	≥26.92
15	Male	≤12.72	12.73–18.65	18.66–23.24	23.25–27.41	≥27.42
15	Female	≤13.65	13.66–19.65	19.66–23.80	23.81–27.89	≥27.90

3.3.2. Physical Fitness Tests and Scoring Criteria

The application implements the five-component battery of the Department of Physical Education, Thailand [2]: body mass index (BMI), sit-and-reach, 30-s modified push-ups, 60-s sit-ups, and the 3-min step test.

The physical fitness performance of students was evaluated based on standardized criteria established by the Department of Physical Education, Thailand. These criteria classify performance into five levels: very good, good, moderate, low, and very low, according to age- and sex-specific benchmarks.

The assessment included four key physical fitness components: flexibility (sit-and-reach), upper-body muscular endurance (30-s modified push-ups), core muscular endurance (60-s sit-ups), and cardiorespiratory endurance (3-min step up-and-down test), as presented in Tables 2–4.

BMI (kg/m^2) was classified into five categories (very thin, thin, proportionate, overweight, and obese) using the age- and sex-specific cut-offs of the Department of Physical Education [2], shown in Table 5.

3.3.3. Expert Evaluation and Manual Reference Classification

Expert evaluation of the application: Nine experts evaluated the application before the field test: four heads of the Health and Physical Education learning area, each with more than 10 years of experience, and five computer and information-technology teachers, each with more than 10 years of experience in software development. The experts evaluated the system by using and testing it and gave feedback, which was used to revise the application before it was tested with teachers at each school.

Manual reference classification: The reference for RQ1 was the manual classification produced by the physical education teachers who administered the tests: they recorded the

raw measurements on paper forms and assigned each student to one of five levels (very low to very good) using the Department of Physical Education criteria [2]. Because the same teachers entered these measurements into the application, the reference is not independent of the application input (see Section 5). In the tables, this reference is labeled "Manual (teachers)" to distinguish it from the expert evaluation described above.

3.3.4. System Usability Scale

The usability of the developed mobile application was assessed using the SUS, a widely recognized and validated instrument for evaluating perceived usability in software systems. The SUS measures key dimensions of usability, including ease of use, efficiency, learnability, and overall user satisfaction, through a set of 10 standardized items administered after system interaction [34–36]. In this study, a total of 36 physical education teachers participated in the usability evaluation. All participants were required to use the developed mobile application in a real-world context prior to completing the SUS questionnaire. Because the evaluation took place after actual use, the responses are based on the teachers' own experience with the system.

Translation procedure: The Thai version of the SUS was used. The English SUS was translated into Thai by a Thai language expert with more than 20 years of experience in English. A foreign expert whose first language is English and who understands Thai then checked the accuracy of the Thai version against the English original. The English and Thai wording of the items is given in Appendix B (Table A1); the Thai wording is not an exact equivalent of the English wording for every item (Section 5). No back-translation or pilot test was carried out, and the respondents were not interviewed.

Scoring: Responses to the ten items (five-point scale; 1 = strongly disagree, 5 = strongly agree) were converted to SUS scores following Brooke [34]. For odd-numbered (positively worded) items the adjusted score was the response minus 1, and for even-numbered (negatively worded) items it was 5 minus the response. The ten adjusted scores (each 0–4) were summed and multiplied by 2.5 to give an overall score from 0 to 100. Item-level means reported in Section 4.2 are expressed on a 0–100 scale (adjusted item score × 25). Scores were interpreted using the acceptability ranges and adjective ratings of Bangor et al. [35] and Sauro [36], with 68 as the benchmark average score.

Reliability: The internal consistency of the SUS in the present sample was estimated with Cronbach's alpha computed on the adjusted item scores (even-numbered items reverse-scored); alpha was also computed separately for the five positively and the five negatively worded items, together with corrected item–total correlations and a bootstrap 95% confidence interval (5000 resamples). Because internal consistency was low, we also examined alpha after the deletion of each item and of the items with non-positive item–total correlations, the response distribution of each item, and the SUS score of the reduced versions rescaled to 0–100. These item analyses were exploratory and were not planned in advance.

3.3.5. User Satisfaction Questionnaire

Teachers' satisfaction was measured with a 20-item questionnaire rated on a five-point Likert scale (results in Section 4.3), covering the usability of the application, the recording of each fitness test, the summarization of results, and data management.

The interpretation of mean scores was based on the following criteria: 4.51–5.00, very high satisfaction; 3.51–4.50, high satisfaction; 2.51–3.50, moderate satisfaction; 1.51–2.50, low satisfaction; and 1.00–1.50, very low satisfaction.

3.4. Procedure

Data were collected through students' physical fitness testing procedures, which were conducted in three main stages, as follows.

(1) Physical Fitness Testing: Physical education teachers administered the physical fitness tests and recorded the results manually using standardized recording forms. The assessment consisted of five components: Body Mass Index (BMI), calculated from weight and height; sit-and-reach test (flexibility); 30-s modified push-ups (upper-body muscular endurance); 60-s sit-ups (core muscular endurance); and 3-min step up-and-down test (cardiorespiratory endurance).
(2) Data Entry into the Mobile Application: Following the testing process, physical education teachers entered the results of all five fitness components into the mobile application. The recorded data were automatically stored in the application's MariaDB/MySQL database.
(3) Data Processing and Output: The application automatically processed the input data, generated evaluation results based on standardized criteria, and stored the information in a structured database. The system also allowed data export in Microsoft Excel format for further statistical analysis.

Evaluation procedure: After using the application in their own teaching context, the 36 teachers completed the SUS and the satisfaction questionnaire online.

Comparison with manual classification: The teachers administered the four performance tests (sit-and-reach, push-ups, sit-ups, and step test), recorded the raw measures manually, and classified each student with the Department of Physical Education criteria [2]. The same raw measures for the 1580 student records were then entered into the application, which applied the age- and sex-specific rules automatically. The records were exported to Microsoft Excel, and the manual and application classifications were compared student by student, followed by a comparison of frequency distributions, means and standard deviations, and Spearman's rank-order correlation.

3.5. Data Analysis

Data were analyzed in three parts. First, descriptive statistics (frequency, percentage, mean, standard deviation, minimum, and maximum) summarized the distribution of fitness classifications and the questionnaire responses. Second, to address RQ1, the agreement between the application's automatic classifications and the teachers' manual classifications of the same measurements was examined with percentage agreement and weighted Cohen's kappa on the five-level classification, with Spearman's rank-order correlation as a supplementary index. These indices verify that the application implements the scoring criteria correctly (computational concordance); they do not constitute independent validation of the measurement procedures or of data entry. Third, to address RQ2 and RQ3, SUS scores were computed as described in Section 3.3.4 and interpreted against established acceptability ranges [35,36], and the satisfaction questionnaire was interpreted using the mean-score ranges given in Section 3.3.5. Internal consistency was estimated with Cronbach's alpha, with $\alpha \geq 0.70$ considered acceptable; the bootstrap confidence interval for alpha was computed in Python 3.12.3 (NumPy 2.4.4, SciPy 1.17.1). All data processing and inferential analyses were performed using IBM SPSS Statistics (version 21.0, IBM Corp., Armonk, NY, USA).

4. Results

4.1. Comparison of Assessment Results

A comparison was conducted between physical fitness assessment results recorded manually by the teachers and those generated by the developed mobile application. The dataset consisted of records from 36 physical education teachers who evaluated the physical fitness test results of 1580 student records (1355 distinct students). Frequencies and percentages in Tables 6–9 refer to these records. The results presented in Table 6 show a comparison of physical fitness assessment outcomes for the sit-and-reach test between manual evaluation and the mobile application. The findings indicate that the distribution of scores across all performance levels is identical between the two methods. The highest proportion of students was classified in the "Very Good" category (41.1%), followed by "Good" (24.4%) and "Moderate" (23.4%). In contrast, the proportions of students in the "Low" and "Very Low" categories were relatively small, accounting for 9.2% and 1.9%, respectively. These results demonstrate a high level of consistency between manual assessments and the mobile application, suggesting that the system classified students' flexibility performance consistently with the manual criteria.

Table 6. A comparative analysis of sit-and-reach test outcomes.

Performance Level	Manual (Teachers)		Mobile Application	
	Records (*n*)	(%)	Records (*n*)	(%)
Very Low	30	1.90	30	1.90
Low	146	9.24	146	9.24
Moderate	369	23.35	369	23.35
Good	385	24.37	385	24.37
Very Good	650	41.14	650	41.14
Total	1580	100.0	1580	100.0

Additionally, the results presented in Table 7 show a comparison of physical fitness assessment outcomes for the 30-s modified push-up test between manual evaluation and the mobile application. The findings indicate that the distribution of scores across all performance levels is identical between the two methods. The highest proportion of students was classified in the "Moderate" category (42.8%), followed by "Good" (22.1%) and "Low" (21.6%). In contrast, the proportions of students in the "Very Low" and "Very Good" categories were relatively small, accounting for 10.1% and 3.4%, respectively.

Table 7. A comparative analysis of the 30-s modified push-up test.

Performance Level	Manual (Teachers)		Mobile Application	
	Records (*n*)	(%)	Records (*n*)	(%)
Very Low	159	10.06	159	10.06
Low	342	21.65	342	21.65
Moderate	677	42.85	677	42.85
Good	349	22.09	349	22.09
Very Good	53	3.35	53	3.35
Total	1580	100.00	1580	100.00

Table 8 presents a comparison of physical fitness assessment outcomes for the 60-s sit-up test between manual evaluation and the mobile application. The findings indicate that the distribution of scores across all performance levels is consistent between the two methods. The highest proportion of students was classified in the "Moderate" category (41.1%), followed by "Good" (26.2%) and "Low" (20.5%). In contrast, the proportions of

students in the "Very Good" and "Very Low" categories were relatively small, accounting for 7.1% and 5.1%, respectively. These results demonstrate a high level of agreement between manual assessments and the mobile application, indicating that the system classified core muscular endurance consistently with the manual criteria.

Table 8. Comparative analysis of the 60-s sit-up test.

Performance Level	Manual (Teachers)		Mobile Application	
	Records (*n*)	(%)	Records (*n*)	(%)
Very Low	80	5.06	80	5.06
Low	324	20.51	324	20.51
Moderate	650	41.14	650	41.14
Good	414	26.20	414	26.20
Very Good	112	7.09	112	7.09
Total	1580	100.00	1580	100.00

Table 9 presents a comparison of physical fitness assessment outcomes for the 3-min step up-and-down test between manual evaluation and the mobile application. The findings indicate that the distribution of scores across all performance levels is identical between the two methods. The highest proportion of students was classified in the "Moderate" category (33.0%), followed by "Very Low" (26.3%) and "Low" (20.4%). In contrast, the proportions of students in the "Good" and "Very Good" categories were relatively smaller, accounting for 15.9% and 4.4%, respectively.

Table 9. Comparative analysis of the 3-min step up-and-down test.

Performance Level	Manual (Teachers)		Mobile Application	
	Records (*n*)	(%)	Records (*n*)	(%)
Very Low	415	26.26	415	26.26
Low	322	20.38	322	20.38
Moderate	522	33.04	522	33.04
Good	252	15.95	252	15.95
Very Good	69	4.37	69	4.37
Total	1580	100.00	1580	100.00

Table 10 presents a comparison of mean scores and standard deviations for physical fitness assessments between manual evaluation and the mobile application across four test components. The results indicate that both methods produced identical mean values and standard deviations for all variables, including sit-and-reach, 30-s modified push-ups, 60-s sit-ups, and the 3-min step up-and-down test.

Table 10. Comparative analysis of physical fitness assessment outcomes.

Variable	Manual (Teachers) (n = 1580)		Application (n = 1580)		r
	Mean	SD	Mean	SD	
Sit-and-Reach	3.94	1.09	3.94	1.09	1.000 *
30-s Modified Push-Ups	2.87	0.98	2.87	0.98	1.000 *
60-s Sit-Ups	3.10	0.97	3.10	0.97	1.000 *
3-min Step Up and Down	2.52	1.16	2.52	1.16	1.000 *

Note: * $p < 0.01$.

Furthermore, the analysis using Spearman's rank-order correlation coefficient revealed perfect positive correlations (r = 1.000, $p < 0.01$) across all variables. This indicates com-

plete agreement between manual assessments and the mobile application. Because the manual and application classifications were derived from the same recorded measurements, this result verifies that the application implements the classification criteria correctly (computational concordance); it does not, by itself, establish the validity of the underlying measurements or the reliability of data entry. The mean scores (±SD) were as follows: sit-and-reach (3.94 ± 1.09), 30-s modified push-ups (2.87 ± 0.98), 60-s sit-ups (3.10 ± 0.97), and 3-min step up-and-down test (2.52 ± 1.16).

To further examine the relationship between the two assessment methods, Spearman's rank-order correlation coefficient was applied. The results revealed that all variables exhibited a perfect positive correlation (r = 1.000), with statistical significance at the 0.01 level ($p < 0.01$). These findings indicate that the assessment results generated by the mobile application are in complete agreement with those obtained from manual evaluation, demonstrating a very high level of consistency in the processing of physical fitness data.

Agreement. The student-by-student comparison covered 1580 student records for each of the four tests (6320 paired classifications). The application and the manual classification agreed for every pair: percentage agreement was 100% for each test, and Cohen's kappa and linearly weighted kappa were 1.00 for each test (no disagreements, so a confidence interval is not informative).

4.2. The System Usability Scale

The usability of the developed mobile application was evaluated using the SUS, a standardized instrument widely used to assess the ease of use, convenience, and overall effectiveness of software systems. The SUS consists of 10 items, with responses collected from users following actual system use [34–36]. In this study, a total of 36 physical education teachers participated in the evaluation by using the application and subsequently completing the SUS questionnaire.

The results presented in Table 11 illustrate the usability evaluation of the developed mobile application for physical education teachers using the SUS, which consists of 10 items. The findings indicate that the mean scores for each item ranged from 75.69 to 98.61, reflecting highly positive user perceptions across all dimensions. When compared with the established SUS benchmark score of 68, all items exceeded the acceptable threshold, indicating a strong level of user acceptance of the system. The item wording in Table 11 is the English source wording; the Thai items administered are given in Table A1 (Appendix B).

Table 11. Evaluating the usability of a mobile application for assessing physical fitness in lower secondary students (aged 13–15 years) using the SUS.

Item	SUS Statement	Score
1	I would like to use this system frequently	93.06
2	I found the system unnecessarily complex	75.69
3	I thought the system was easy to use	87.50
4	I think I would need technical support to use this system	85.42
5	I found the system's functions well integrated	98.61
6	I thought there was too much inconsistency in the system	88.89
7	I think most people would learn to use this system quickly	88.19
8	I found the system very cumbersome to use	79.17
9	I felt confident using the system	89.58
10	I needed to learn a lot before I could get going	84.72

A detailed item-level analysis revealed that Item 1 (Mean = 93.06) suggests that users have a highly positive attitude toward the system and are likely to adopt it in practice. Item 2 (Mean = 75.69), which recorded the lowest score among all items, still falls within a high range, indicating that some users may perceive a degree of complexity in certain system

functions; however, this complexity remains within acceptable limits. Item 3 (Mean = 87.50) demonstrates that the system is easy to use and can be learned quickly. Item 4 (Mean = 85.42) indicates that most users are able to operate the system independently without requiring extensive technical support.

Item 5 (Mean = 98.61), the highest score observed, reflects a very high level of user satisfaction with the system's functionality and effectiveness. Item 6 (Mean = 88.89) suggests that the system operates with high stability and continuity, minimizing errors during use. Item 7 (Mean = 88.19) indicates that the system has good learnability, enabling new users to become proficient within a short period. Item 8 (Mean = 79.17) implies that some users may still perceive minor complexity; however, overall usability remains high. Item 9 (Mean = 89.58) reflects strong user confidence in the system, which is a critical indicator of software quality. Finally, Item 10 (Mean = 84.72) suggests that while some initial learning may be required, the system is not overly complex or difficult to use.

Overall, the aggregated SUS score indicates a very high level of usability and falls within the acceptable range. The results suggest a strong likelihood of real-world adoption in physical education settings. These findings clearly demonstrate that the developed application possesses high usability quality in terms of ease of use, convenience, and system efficiency, thereby supporting its effective implementation in educational contexts.

Internal consistency: Cronbach's alpha for the ten SUS items was 0.42 (95% bootstrap CI −0.07 to 0.62), below the conventional threshold of 0.70. The five positively worded items had $\alpha = 0.72$, whereas the five negatively worded items had $\alpha = 0.16$, and the corrected item–total correlations of items 2, 4, and 10 were negative or close to zero (−0.22, −0.10, and −0.04). Item scores were high with little variance, particularly for item 5 (SD = 0.23 on the 0–4 scale). The mean SUS score should therefore be interpreted with caution.

Item analysis: Table 12 shows that the low consistency is concentrated in the negatively worded items 2, 4, and 10, whose corrected item-total correlations were negative or near zero; deleting any one of them raised alpha to 0.56 (item 2), 0.50 (item 4), and 0.49 (item 10). Deleting items 2, 4, and 10 together gave a seven-item set with $\alpha = 0.74$ (95% bootstrap CI 0.53 to 0.82). The responses were also restricted in range: no teacher rated any negatively worded statement 4 or 5, and 34 of 36 teachers gave the highest rating to item 5. The SUS score rescaled to 0–100 was 89.3 for the seven-item set and 91.4 for the five positively worded items alone ($\alpha = 0.72$), compared with 87.1 for the full scale, so all versions exceeded the benchmark of 68.

Table 12. Item analysis of the Thai System Usability Scale (n = 36). Responses are the numbers of teachers choosing 1 (strongly disagree) to 5 (strongly agree) on the original response scale; corrected item-total correlations and alpha if the item is deleted were computed on the adjusted (reverse-scored) item scores.

Item	Tone	Responses (1/2/3/4/5)	Corrected Item–Total r	α If Item Deleted
1	Positive	0/0/1/8/27	0.46	0.29
2	Negative	10/17/9/0/0	−0.22	0.56
3	Positive	0/0/3/12/21	0.44	0.27
4	Negative	18/15/3/0/0	−0.10	0.50
5	Positive	0/0/0/2/34	0.09	0.42
6	Negative	20/16/0/0/0	0.33	0.34
7	Positive	0/0/1/15/20	0.24	0.37
8	Negative	10/22/4/0/0	0.47	0.27
9	Positive	0/0/0/15/21	0.34	0.34
10	Negative	19/12/5/0/0	−0.04	0.49
All 10 items				0.42

The usability evaluation of the mobile application was estimated using the SUS among physical education teachers. The findings indicate that the overall SUS score achieved a mean of 87.08, with a minimum score of 72.50 and a maximum score of 97.50, out of a total possible score of 100. The SUS score is depicted in Figure 5.

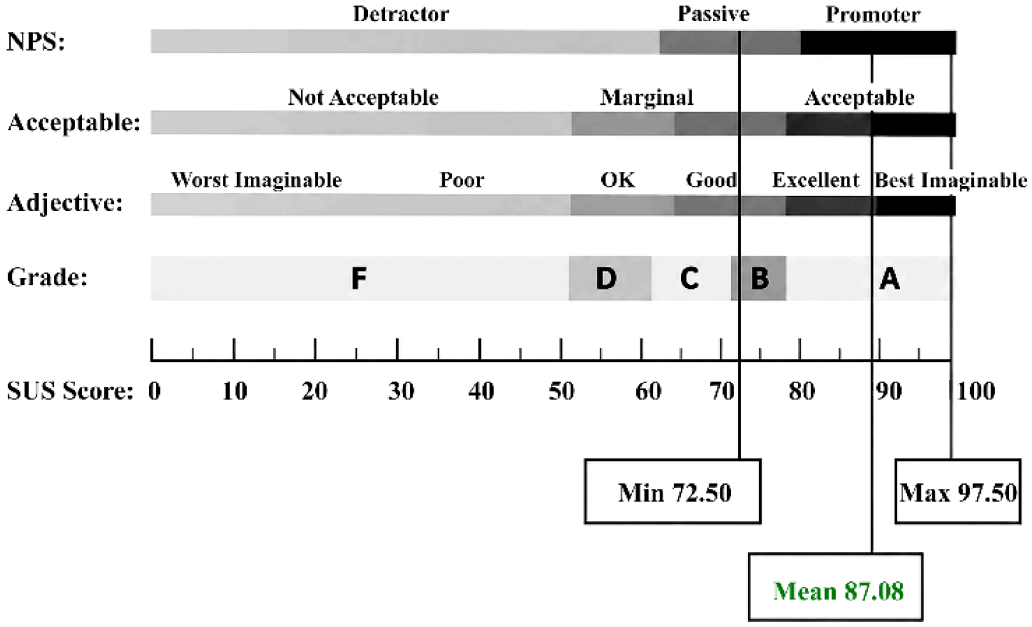


Figure 5. Summary of the SUS evaluation results for the mobile application used by physical education teachers to assess the physical fitness of lower secondary school students (aged 13–15 years) [36].

Based on established SUS interpretation criteria, the mean score of 87.08 falls within the 80–90 range, corresponding to an "Excellent" rating and a Grade A classification. Furthermore, the score lies within the "Acceptable" usability range, indicating a high level of user acceptance. In addition, according to the Net Promoter Score (NPS) framework, the system is categorized within the "Promoter" group, suggesting that users are highly likely to recommend the application to others.

Notably, the minimum score of 72.50, although the lowest among participants, still falls within the "Good" category and remains within the acceptable usability range, indicating that all users maintained a positive perception of the system. Conversely, the maximum score of 97.50 corresponds to the "Best Imaginable" category, reflecting exceptionally high user satisfaction.

These findings suggest that variations in technological proficiency among physical education teachers did not present a significant barrier to system use. Overall, users were able to effectively understand and operate the mobile application, demonstrating its high usability, accessibility, and potential for widespread adoption in educational settings.

4.3. Assessment of User Satisfaction Among Physical Education Teachers

The evaluation of user satisfaction among physical education teachers regarding the use of the mobile application for assessing the physical fitness of lower secondary school students (aged 13–15 years) was conducted using a five-point Likert-scale questionnaire. The sample consisted of 36 physical education teachers who completed the evaluation after using the developed application.

The assessment aimed to examine user satisfaction across several dimensions, including system usability, data recording, summarization of physical fitness test results, and the overall efficiency of the application in facilitating data management.

The results presented in Table 13 illustrate the level of user satisfaction among physical education teachers regarding the use of the mobile application for assessing the physical fitness of lower secondary school students (aged 13–15 years). The findings indicate that the overall satisfaction of the 36 respondents was at a very high level (Mean = 4.51, SD = 0.65).

Table 13. User satisfaction of physical education teachers with the mobile application for assessing the physical fitness of lower secondary school students (aged 13–15 years).

Item	Evaluation Criteria	Mean	S.D.	Interpretation
1	Application installation on mobile devices	4.44	0.65	High
2	Application startup and accessibility	4.44	0.69	High
3	Menu layout and interface organization	4.41	0.65	High
4	Recording weight and height for BMI calculation	4.58	0.60	Very High
5	Recording sit-and-reach test data	4.63	0.68	Very High
6	Recording 30-s modified push-ups	4.47	0.74	High
7	Recording 60-s sit-ups	4.66	0.59	Very High
8	Recording 3-min step test	4.63	0.59	Very High
9	Male fitness standard criteria	4.52	0.61	Very High
10	Female fitness standard criteria	4.47	0.70	High
11	BMI result reporting	4.61	0.60	Very High
12	Sit-and-reach result reporting	4.58	0.60	Very High
13	Push-up result reporting	4.63	0.68	Very High
14	Sit-up result reporting	4.55	0.69	High
15	Step test result reporting	4.58	0.60	Very High
16	Ease of recording using the application	4.30	0.71	High
17	Time-saving capability	4.36	0.72	High
18	Reduction in recording errors	4.52	0.70	Very High
19	Improvement of data accuracy	4.44	0.65	High
20	Database storage capability	4.44	0.69	High

A detailed analysis of individual items revealed that the highest level of satisfaction was associated with the recording function for the 60-s sit-up test (Mean = 4.66, SD = 0.59), reflecting users' strong approval of the system's functionality in capturing test data. In contrast, the lowest level of satisfaction was related to the convenience of using the program for recording data (Mean = 4.30, SD = 0.71). However, this score still falls within the high satisfaction level, indicating that users generally perceive the system as effective and practical for data recording. Overall, these findings suggest that the developed mobile application is well received by users and demonstrates strong usability and practical value in supporting physical fitness assessment in educational settings.

5. Discussion

This study aimed to develop and evaluate a mobile application to support physical education teachers in assessing the physical fitness of lower secondary school students aged 13–15 years. The application was developed using a Research and Development (R&D) approach based on the System Development Life Cycle (SDLC), including the planning, design, development, testing, and implementation phases, to ensure that the system met user requirements and achieved high functional performance [37–39].

The study involved 36 physical education teachers from secondary schools in Mueang District, Chiang Mai Province. Most participants were male (75%) and aged 20–40 years, a demographic generally familiar with digital technologies and more likely to adopt mobile applications effectively [24,25,39–41].

The results demonstrated that the application recorded physical fitness data, calculated Body Mass Index (BMI), and generated assessment results based on the standardized criteria of the Department of Physical Education, Thailand. Comparison with manual

assessments showed identical mean scores across all fitness components, indicating that the application correctly implements the classification criteria. Whether the application reduces recording errors and processing time compared with manual methods, as information-systems principles would suggest [39–41], was not quantified in this study and should be tested in future work.

The usability evaluation produced a mean System Usability Scale (SUS) score of 87.08, indicating excellent usability and strong user acceptance [34,35]. Participants also reported very high satisfaction with the system, particularly regarding usability, data accuracy, processing speed, and database management, consistent with previous studies demonstrating the benefits of digital technologies for physical fitness assessment and data management [39–41].

Source of the low internal consistency. The data narrow the possible sources but cannot identify a single one. First, the problem lies in three negatively worded items (2, 4, and 10), whereas the other negatively worded items (6 and 8) correlated positively with the rest of the scale ($r = 0.33$ and 0.47) and the five positively worded items were acceptably consistent ($\alpha = 0.72$). Second, responses were strongly restricted in range (Table 12), which lowers inter-item correlations mechanically, especially with 36 respondents. Third, we found no sign of the mistake pattern described for mixed-tone SUS data, in which respondents accidentally agree with negative statements [42]: no teacher rated a negative statement 4 or 5, and no teacher gave the same answer to all items. Fourth, items 2 and 4 correlated weakly and negatively with positively worded items (for example, $r = -0.28$ between items 1 and 2); teachers who rated the system most favorably were somewhat more likely to answer "neutral" than "strongly disagree" to the statement that it was unnecessarily complex, which may reflect real perceived complexity in some functions rather than a measurement error. Mixed-tone questionnaires can also form separate positive-tone and negative-tone factors [43], and the sums of the two sets of items were uncorrelated in this sample ($r = 0.00$). A problem of translation or cultural adaptation, such as the Thai rendering of negation or of the statements about complexity, technical support, and learning, cannot be excluded and the procedure described in Section 3.3.4 included no back-translation or pilot test with teachers and no interviews with the respondents. In addition, the Thai wording of several items differs in meaning from the English wording (see below). Testing these explanations requires cognitive interviews with Thai teachers, or a comparison with an all-positively worded version [42].

Thai and English wording. Table A1 shows that the Thai rendering departs from the English wording in several items. Item 5 ("functions well integrated") was rendered as "many functions work well", so it no longer refers to integration; item 2 ("unnecessarily complex") was rendered as "the system should not be this complex"; item 10 ("learn a lot before I could get going") was rendered as wanting to practice before starting, without "a lot"; item 4 states that technical or IT help was needed before the system could be used; and items 4 and 8 contain alternative terms separated by a slash. Items 2, 4, and 10 are the items with non-positive item–total correlations, so a translation effect on these items is plausible, and the Thai version should not be regarded as strictly equivalent to the standard SUS. This comparison is our reading of the Thai text and has not been checked with the translators.

Effect of removing items: Removing items 2, 4, and 10 raised alpha from 0.42 to 0.74, but we kept the full ten-item scale for the main analysis because the deletion is data-driven and exploratory in a small sample, and because a reduced score cannot be compared with published SUS benchmarks. The conclusion does not depend on this choice: the rescaled mean was 89.3 for the seven-item set and 91.4 for the positively worded items alone, against 87.1 for the full scale, all above the benchmark of 68. The SUS result is therefore best read as

an indication of favorable perceived usability, and the reliability of the Thai version needs to be established in a larger sample.

Overall, the findings provide empirical evidence that the developed mobile application is a usable, user-centered tool that applies the national classification criteria consistently in school-based physical fitness assessment and that teachers rated highly. Its potential to enhance the efficiency and standardization of fitness evaluation supports digital transformation in physical education, but this requires confirmation with independent, quantitative evidence. Future research should investigate large-scale implementation, long-term impacts on student health outcomes, and the integration of advanced technologies, such as data analytics and artificial intelligence, to further improve system capabilities.

This study has several limitations that should be acknowledged. First, data collection was restricted to physical education teachers and students within Mueang District, Chiang Mai Province, which may limit the generalizability of the findings to other regions or educational contexts with different resources and infrastructure. Second, although the sample size of 36 teachers was statistically determined using Yamane's formula, this remains a relatively modest sample, which may limit statistical power for subgroup analyses and the detection of smaller effects. Third, the developed application currently assesses only five physical fitness components (BMI, flexibility, muscular strength and endurance, and cardiorespiratory endurance) and does not yet incorporate skill-related components such as agility, balance, and reaction time, which would allow for a more comprehensive assessment of students' physical fitness. Fourth, the comparison between application-generated results and manual assessments was conducted without blinding, which may introduce potential rater bias; future studies should consider blinded or independent comparisons to further strengthen validity evidence. Fifth, the manual and application classifications were derived from the same recorded measurements and data were still entered manually into the application, so the study verifies the application's scoring logic but does not independently validate the measurements or the accuracy of data entry. Sixth, no data on recording time or error rates were collected for either the conventional manual process or the application, so the benefits for efficiency and error reduction could not be quantified, and the usability results come from a single group of teachers without a comparison condition. Seventh, this study employed a cross-sectional design conducted over a limited period and therefore does not provide evidence regarding the long-term reliability, sustained usability, or durability of the application with continued use. Eighth, the Thai version of the SUS was prepared through expert translation and verification, but its internal consistency in this sample was low ($\alpha = 0.42$) and its other psychometric properties were not evaluated, the Thai wording of several items differs in meaning from the English wording (Table A1), and the source of the low consistency could not be identified because no cognitive interviews were conducted and the item analysis is exploratory, so the SUS score should be interpreted with caution. Ninth, the 1580 records come from 36 teachers who partly entered the same student data sets (1355 distinct students), so the records are not independent observations of different students, and the frequencies in Tables 6–9 should not be read as estimates of the fitness of a student population. Finally, successful implementation of the application depends on access to computers or mobile devices with a web browser and stable internet connectivity, which may pose a barrier to adoption in schools with limited technological resources. Future research should address these limitations by expanding the sample to a more geographically diverse population, incorporating additional skill-related fitness components, employing blinded assessment procedures, and evaluating the application's long-term usability and impact across varied educational settings.

6. Conclusions

This study developed and evaluated a mobile application to support physical education teachers in assessing the physical fitness of lower secondary school students aged 13–15 years using the System Development Life Cycle (SDLC). The application's automatic classifications were identical to manual classifications of the same measurements across all fitness tests analyzed, indicating that it implements the national scoring criteria correctly; whether it reduces processing time and recording errors remains to be quantified. Usability evaluation yielded a high System Usability Scale (SUS) score (M = 87.08), although internal consistency of the Thai SUS in this sample was low ($\alpha = 0.42$), while teachers reported high satisfaction with its usability, accuracy, speed, and practical applicability. These findings suggest that the application is a usable, user-centered tool for physical fitness assessment with promising potential for implementation in school-based physical education.

Author Contributions: Conceptualization, N.F., T.K. and J.L.; methodology, N.F. and T.K.; software, N.F.; validation, N.F., J.L. and T.K.; formal analysis, N.F., J.L. and T.K.; investigation, N.F.; resources, N.F.; data curation, N.F.; writing—original draft preparation, N.F. and T.K.; writing—review and editing, N.F., J.L. and T.K.; supervision, T.K. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by Chiang Mai University, Chiang Mai, Thailand.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Research Ethics Committee of the Multidisciplinary and Interdisciplinary College, Chiang Mai University (protocol code MIdSEC-68EX-003, approved 12 January 2026).

Informed Consent Statement: Written informed consent was obtained from all participants involved in the study. A consent form outlining the study procedures, potential risks, and the voluntary nature of participation was distributed to each participant prior to enrollment, and each participant signed the form before taking part in the study.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to acknowledge the College of Arts, Media, and Technology, Chiang Mai University, and the Multidisciplinary and Interdisciplinary School, Chiang Mai University, Thailand, for their facilities support. The authors would also like to express their special thanks to Sonchai Jaisan for advising on the development of the program and also Chaiinarong Sinpoo for translating the contents of this paper. In addition, the authors would like to thank the participants for their valuable contributions to the research.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BMI	Body mass index
HIIT	High-intensity interval training
NPS	Net Promoter Score
PDO	PHP Data Objects
PE	Physical education
R&D	Research and development
SDLC	System development life cycle
SUS	System Usability Scale

Appendix A

This appendix presents representative screenshots of the data management interface of the developed mobile application, illustrating how physical fitness test records for each of the five components are displayed, entered, and managed by physical education teachers. For each test, the application provides two complementary views: a records view, which lists all measured values alongside the student's gender and age group for reference and verification, and a management view, which allows teachers to search, edit, import, and export test records as presented in Figures A1–A7.

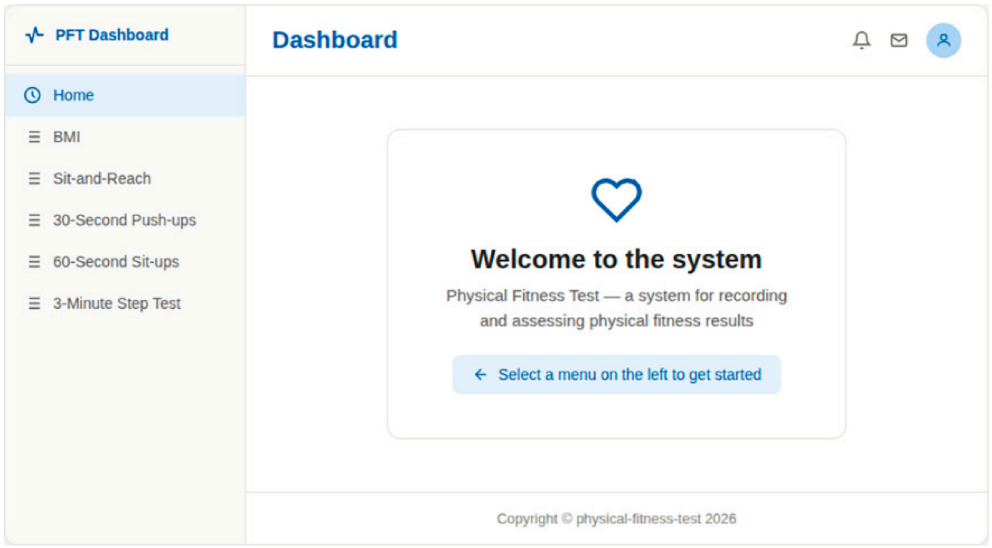


Figure A1. Main Menu of the Physical Fitness Test application (PFT Dashboard).

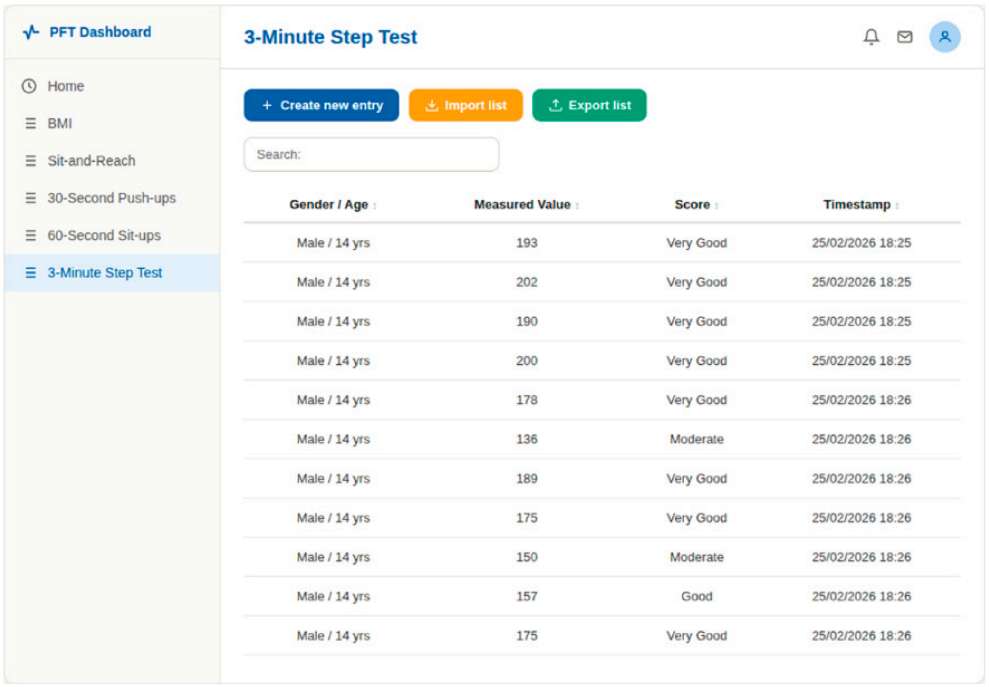


Figure A2. Records view of the 3-min step test, showing the measured value, computed score classification, and timestamp for each entry.

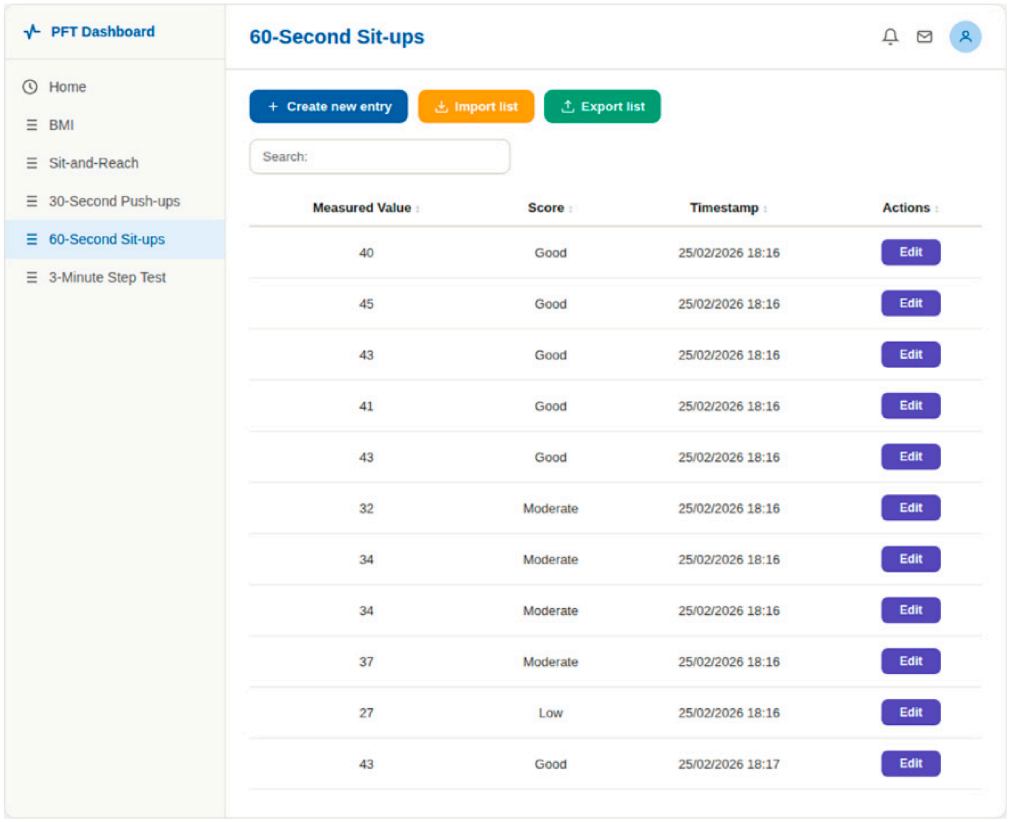


Figure A3. Management view of the 60-s sit-up test, showing the measured value, score classification, timestamp, and an edit function for correcting individual records.

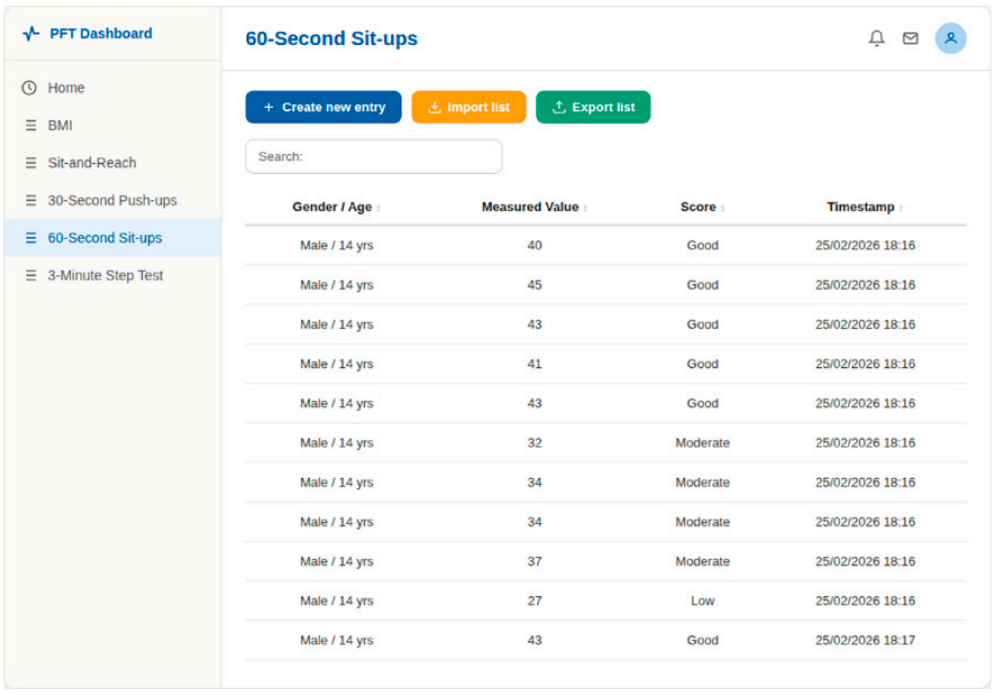


Figure A4. Records view of the 60-s sit-up test, showing the same data alongside each student's gender and age group.

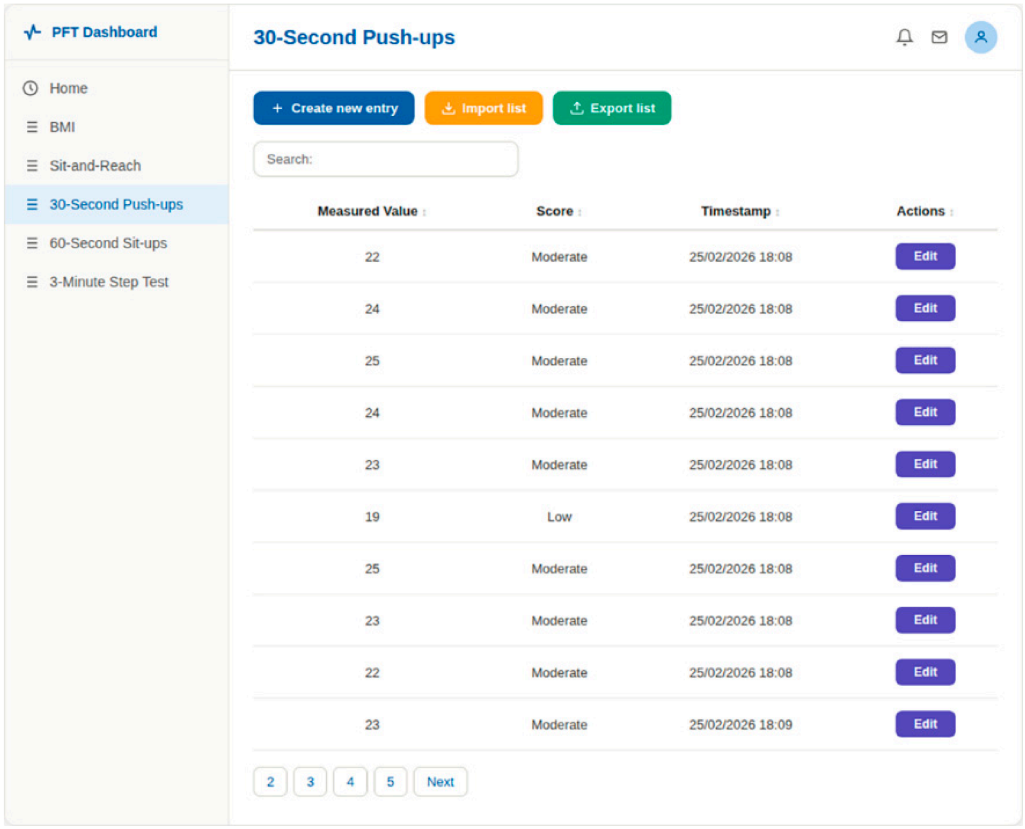


Figure A5. Records view of the 30-s push-up test, showing the measured value, score classification, and timestamp alongside each student's gender and age group.

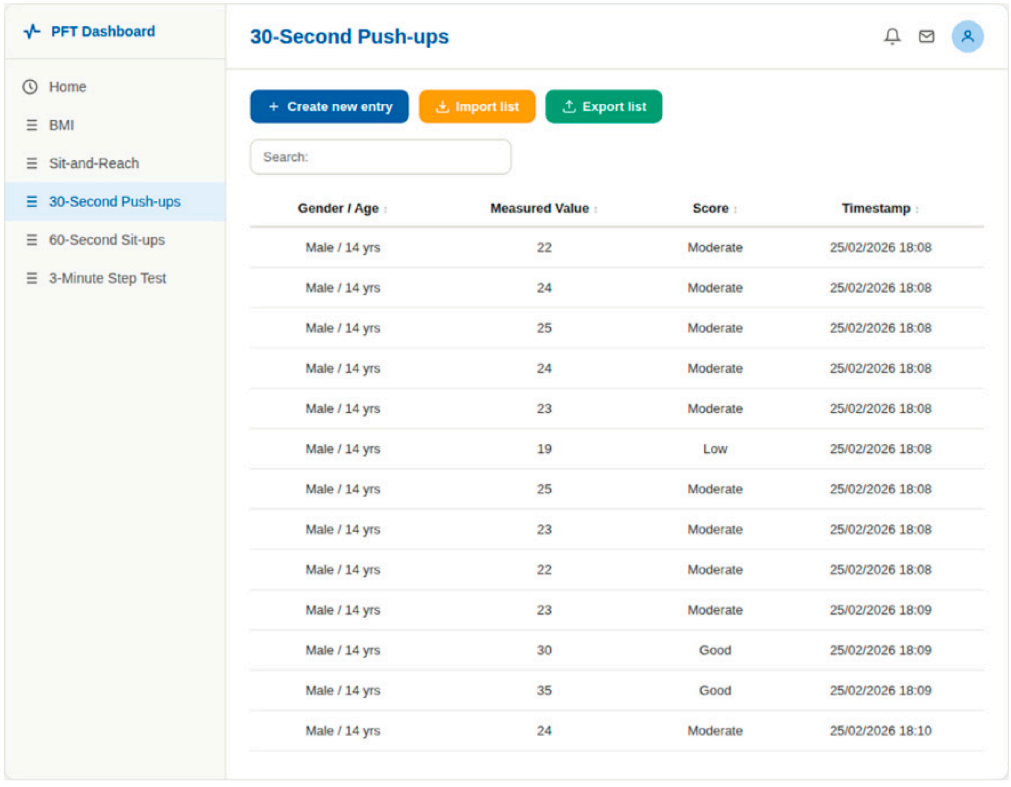


Figure A6. Management view of the 30-s push-up test, showing the edit function and paginated navigation across multiple pages of records.

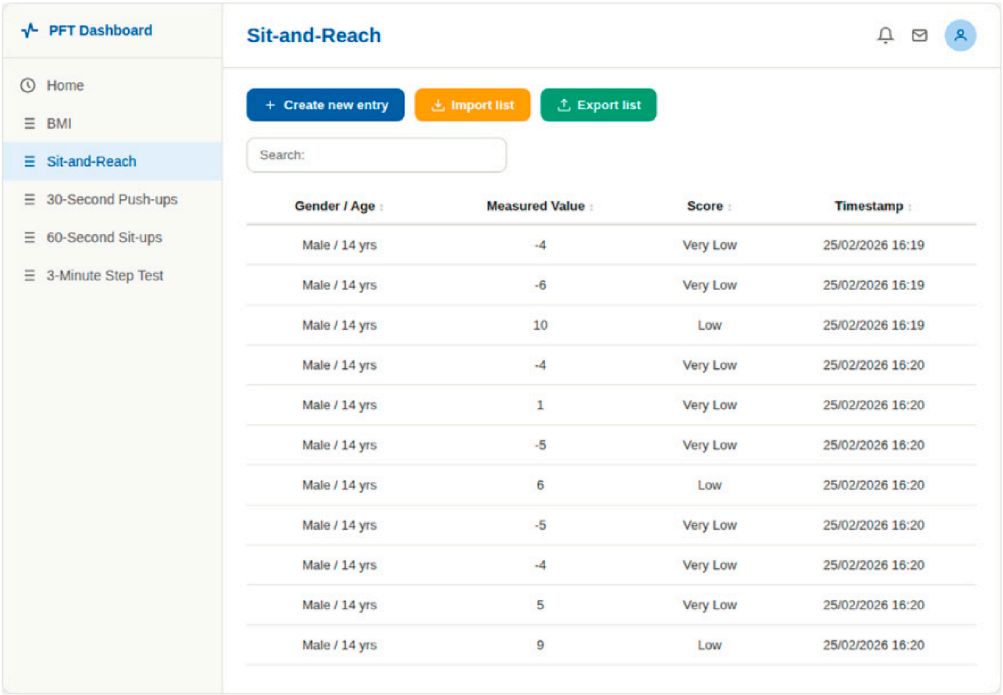


Figure A7. Records view of the sit-and-reach test, showing the measured value (in centimeters), score classification, and timestamp alongside each student's gender and age group.

Appendix B

Table A1 gives the English wording of the ten SUS items used in this study and the Thai version administered to the teachers (translation procedure in Section 3.3.4). The English wording is the source wording reported in Table 11; the Thai items are given as administered and differ in meaning from the English items for some items (Section 5).

Table A1. System Usability Scale items in English and in the Thai version administered to physical education teachers.

Item	English Wording	Thai Version Administered
1	I would like to use this system frequently	ฉันอยากใช้ระบบนี้บ่อย ๆ
2	I found the system unnecessarily complex	ฉันคิดว่า ระบบไม่ควรซับซ้อนขนาดนี้
3	I thought the system was easy to use	ฉันคิดว่าระบบนี้ใช้ง่าย
4	I think I would need technical support to use this system	ฉันคิดว่าต้องมีคนฝั่งเทคนิค/ไอทีเข้ามาช่วย ฉันถึงจะใช้งานระบบนี้ได้
5	I found the system's functions well integrated	ฉันพบว่ามีหลายฟังก์ชันที่ทำงานได้ดี
6	I thought there was too much inconsistency in the system	ฉันคิดว่าระบบไม่ค่อยมีความสม่ำเสมอ
7	I think most people would learn to use this system quickly	ฉันคิดว่าคนอื่น ๆ น่าจะเข้าใจวิธีใช้ระบบนี้ได้เร็วเหมือนกัน
8	I found the system very cumbersome to use	ฉันพบว่าการใช้ระบบยุ่งยาก/ซับซ้อนมาก ๆ
9	I felt confident using the system	ฉันรู้สึกมั่นใจตอนใช้งาน
10	I needed to learn a lot before I could get going	ฉันต้องการฝึกใช้งานก่อนถึงจะเริ่มใช้งานระบบนี้ได้

References

1. Bergen, T.; Richards, J.; Kira, G.; Kim, A.H.M.; Signal, L.; Mizdrak, A. The effectiveness of secondary-school-based interventions on the future physical activity of adolescents in Aotearoa New Zealand: A modelling study. *Int. J. Behav. Nutr. Phys. Act.* **2024**, *21*, 114. [CrossRef] [PubMed]
2. Department of Physical Education. *Physical Fitness Tests and Standards for Students Aged 13–18 Years*; Ministry of Tourism and Sports: Bangkok, Thailand, 2019. Available online: https://prgroup.hss.moph.go.th/attachments/article/709/ebook02.pdf (accessed on 10 May 2026).

3. Bull, F.C.; Al-Ansari, S.S.; Biddle, S.; Borodulin, K.; Buman, M.P.; Cardon, G.; Carty, C.; Chaput, J.P.; Chastin, S.; Chou, R.; et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br. J. Sports Med.* **2020**, *54*, 1451–1462. [CrossRef] [PubMed]
4. Department of Physical Education. *Physical Fitness Test Storage and Reporting Program*; Ministry of Tourism and Sports: Bangkok, Thailand, 2021. Available online: https://sportsci.dpe.go.th/pages/index/ (accessed on 1 November 2026).
5. den Uijl, I.; Ter Hoeve, N.; Sunamura, M.; Lenzen, M.J.; Braakhuis, H.E.M.; Stam, H.J.; Boersma, E.; van den Berg-Emons, R.J.G. Physical Activity and Sedentary Behavior in Cardiac Rehabilitation: Does Body Mass Index Matter? *Phys. Ther.* **2021**, *101*, pzab142. [CrossRef] [PubMed]
6. Samahito, S. Physical Fitness and Health Education. Bachelor's Thesis, Bachelor of Science (Sports Science), Faculty of Education Kasetsart University, Bangkok, Thailand, 1998.
7. Riebe, D.; Ehrman, J.K.; Liguori, G.; Magal, M. *ACSM's Guidelines for Exercise Testing and Prescription*, 10th ed.; Wolters Kluwer: Philadelphia, PA, USA, 2018.
8. World Health Organization. *WHO Guidelines on Physical Activity and Sedentary Behaviour*; World Health Organization: Geneva, Switzerland, 2020. Available online: https://www.who.int/publications/i/item/9789240015128 (accessed on 10 May 2026).
9. World Health Organization. *Global Status Report on Physical Activity 2022*; World Health Organization: Geneva, Switzerland, 2022. Available online: https://www.who.int/publications/i/item/9789240059153 (accessed on 26 August 2026).
10. Guthold, R.; Stevens, G.A.; Riley, L.M.; Bull, F.C. Global trends in insufficient physical activity among adolescents: A pooled analysis of 298 population-based surveys with 1.6 million participants. *Lancet Child Adolesc. Health* **2020**, *4*, 23–35. [CrossRef] [PubMed]
11. Kang, P.J.; Kuo, H.C. Optimizing adolescent health: Investigating physical activity's impact on fitness, working memory, and academic performance. *Cogn. Dev.* **2024**, *71*, 101478. [CrossRef]
12. Fu, Q.; Li, L.; Li, Q.; Wang, J. The effects of physical activity on the mental health of typically developing children and adolescents: A systematic review and meta-analysis. *BMC Public Health* **2025**, *25*, 1514. [CrossRef] [PubMed]
13. Donnelly, J.E.; Hillman, C.H.; Castelli, D.; Etnier, J.L.; Lee, S.; Tomporowski, P.; Lambourne, K.; Szabo-Reed, A.N. Physical activity, fitness, cognitive function, and academic achievement in children: A systematic review. *Med. Sci. Sports Exerc.* **2016**, *48*, 1197–1222. [CrossRef] [PubMed]
14. Mitić, P.; Jovanović, R.; Stojanović, N.; Barišić, V.; Trajković, N. Enhancing adolescent physical fitness and well-being: A school-based high-intensity interval training program. *J. Funct. Morphol. Kinesiol.* **2024**, *9*, 279. [CrossRef] [PubMed]
15. Gasser, M.; Nadenbousch, A.M.; Egger, F.; Kamer, M.; Valkanover, S.; Schmidt, M. Increasing adolescents' physical activity levels through a comprehensive school-based physical activity program: Study protocol of the cluster randomized controlled trial Active School. *BMC Pediatr.* **2024**, *24*, 561. [CrossRef] [PubMed]
16. He, H.; Yang, Y.; Sun, J.; Wang, F.; Zhang, W.; Zhu, F. Effects of school-based physical activity on academic achievement in children and adolescents: A systematic review and meta-analysis. *Front. Public Health* **2025**, *13*, 1651883. [CrossRef] [PubMed]
17. Pérez-Ramírez, J.A.; Santos, M.P.; Mota, J.; González-Fernández, F.T.; Villa-González, E. Enhancing adolescent health: The role of strength and endurance school-based HIIT interventions in physical fitness and cognitive development. *Front. Psychol.* **2025**, *16*, 1568129. [CrossRef] [PubMed]
18. Silva, A.M.; Graça, A.; MacPhail, A.; Batista, P. School-based physical activity programmes: An integrative review of health-related research. *Stud. Sport.* **2025**, *19*, 102–118. [CrossRef]
19. Wang, H.; Pang, J.; Yang, X.; Jia, Y.; Huang, X.; Yu, L.; Hou, X. School-based environment and physical activity in adolescents: A systematic review and meta-analysis. *Prev. Med.* **2025**, *191*, 108221. [CrossRef] [PubMed]
20. Duncombe, S.L.; Barker, A.R.; Bond, B.; Earle, R.; Varley-Campbell, J.; Vlachopoulos, D.; Walker, J.L.; Weston, K.L.; Stylianou, M. School-based high-intensity interval training programs in children and adolescents: A systematic review and meta-analysis. *PLoS ONE* **2022**, *17*, e0266427. [CrossRef] [PubMed]
21. Mateo-Orcajada, A.; Vaquero-Cristóbal, R.; Mota, J.; Abenza-Cano, L. Physical activity, body composition, and fitness variables in adolescents after periods of mandatory, promoted or nonmandatory, nonpromoted use of step tracker mobile apps: Randomized controlled trial. *JMIR mHealth uHealth* **2024**, *12*, e51206. [CrossRef] [PubMed]
22. Bano, M.; Zowghi, D. Users' involvement in mobile app development: A systematic mapping study. *Inf. Softw. Technol.* **2015**, *65*, 80–99.
23. Ngamwannakorn, C. *Development of a Learning Application on Mobile Device in English for Information Officer*; Report No. 010–2562; Faculty of Humanities and Social Sciences, Yala Rajabhat University: Yala, Thailand, 2019.
24. Direito, A.; Jiang, Y.; Delbaere, K.; Maddison, R.; Thomas, K.; Foley, B. Apps for IMproving FITness and increasing physical activity among young people: The AIMFIT pragmatic randomized controlled trial. *J. Med. Internet Res.* **2015**, *17*, e210. [CrossRef] [PubMed]

25. Recio-Rodriguez, J.I.; Agudo-Conde, C.; Martin-Cantera, C.; González-Viejo, M.N.; Fernandez-Alonso, M.D.; Arietaleanizbeaskoa, M.S.; Schmolling-Guinovart, Y.; Maderuelo-Fernandez, J.A.; Rodriguez-Sanchez, E.; Gomez-Marcos, M.A.; et al. Short-term effectiveness of a mobile phone app for increasing physical activity and adherence to the Mediterranean diet in primary care: A randomized controlled trial (EVIDENT II Study). *J. Med. Internet Res.* **2016**, *18*, e331. [CrossRef] [PubMed]
26. Ruiz, J.R.; Castro-Piñero, J.; España-Romero, V.; Artero, E.G.; Ortega, F.B.; Cuenca, M.M.; Jimenez-Pavón, D.; Chillón, P.; Girela-Rejón, M.J.; Mora, J.; et al. Field-based fitness assessment in young people: The ALPHA health-related fitness test battery for children and adolescents. *Br. J. Sports Med.* **2011**, *45*, 518–524. [CrossRef] [PubMed]
27. Mateo-Orcajada, A.; Abenza-Cano, L.; Albaladejo-Saura, M.D.; Vaquero-Cristóbal, R. Mandatory after-school use of step tracker apps improves physical activity, body composition and fitness of adolescents. *Educ. Inf. Technol.* **2023**, *28*, 10235–10266. [CrossRef] [PubMed]
28. Tohănean, D.I.; Vulpe, A.M.; Mijaica, R.; Alexe, D.I. Embedding digital technologies (AI and ICT) into physical education: A systematic review of innovations, pedagogical impact, and challenges. *Appl. Sci.* **2025**, *15*, 9826. [CrossRef]
29. Zhong, Q.; Jiang, J.; Bai, W.; Yin, Z.; Liao, Z.; Zhong, X. Application of digital-intelligent technologies in physical education: A systematic review. *Front. Public Health* **2025**, *13*, 1626603. [CrossRef] [PubMed]
30. Parmenter, B.; Burley, C.; Stewart, C.; Whife, J.; Champion, K.; Osman, B.; Newton, N.; Green, O.; Wescott, A.B.; Gardner, L.A.; et al. Measurement properties of smartphone approaches to assess physical activity in healthy young people: Systematic review. *JMIR mHealth uHealth* **2022**, *10*, e39085. [CrossRef] [PubMed]
31. Bogataj, Š.; Pajek, M.; Hadžić, V.; Andrašić, S.; Padulo, J.; Trajković, N. Validity, reliability, and usefulness of My Jump 2 app for measuring vertical jump in primary school children. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3708. [CrossRef] [PubMed]
32. Sulla-Torres, J.; Santos-Pamo, B.; Cárdenas-Rodríguez, F.; Angulo-Osorio, J.; Gómez-Campos, R.; Cossio-Bolaños, M. Multiplatform computer vision system to support physical fitness assessments in schoolchildren. *Appl. Sci.* **2024**, *14*, 7140. [CrossRef]
33. Joskow, J. Reviewed Work: Statistics, an Introductory Analysis. by Taro Yamane. *J. Am. Stat. Assoc.* **1965**, *60*, 678–680.
34. Brooke, J. SUS: A quick and dirty usability scale. In *Usability Evaluation in Industry*; Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L., Eds.; Taylor & Francis: London, UK, 1996; pp. 189–194. [CrossRef]
35. Bangor, A.; Kortum, P.T.; Miller, J.T. An empirical evaluation of the System Usability Scale. *Int. J. Hum.-Comput. Interact.* **2008**, *24*, 574–594. [CrossRef]
36. Sauro, J. 5 Ways to Interpret a SUS Score. MeasuringU. 2018. Available online: https://measuringu.com/interpret-sus-score/ (accessed on 3 March 2026).
37. AlHazme, R.H.; Rana, A.M.; De Lucca, M. Development and implementation of a clinical and business intelligence system for the Florida health data warehouse. *Online J. Public Health Inform.* **2014**, *6*, e182. [CrossRef] [PubMed]
38. Aggelidis, V.P.; Chatzoglou, P.D. Hospital information systems: Measuring end user computing satisfaction (EUCS). *J. Biomed. Inform.* **2012**, *45*, 566–579. [CrossRef] [PubMed]
39. Semsem, K.; Martin, J.T. Development of a mobile application for physical fitness testing. *Int. J. Hum. Mov. Sports Sci.* **2022**, *10*, 1126–1133. [CrossRef]
40. Middelweerd, A.; Mollee, J.S.; van der Wal, C.N.; Brug, J.; Te Velde, S.J. Apps to promote physical activity among adolescents: A review and content analysis. *Int. J. Behav. Nutr. Phys. Act.* **2014**, *11*, 97. [CrossRef] [PubMed]
41. Direito, A.; Dale, L.P.; Shields, E.; Wedderburn, R.; Whittaker, R.; Maddison, R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? A systematic review. *BMC Public Health* **2014**, *14*, 646. [CrossRef] [PubMed]
42. Sauro, J.; Lewis, J.R. When designing usability questionnaires, does it hurt to be positive? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; pp. 2215–2224.
43. Lewis, J.R. The System Usability Scale: Past, present, and future. *Int. J. Hum.-Comput. Interact.* **2018**, *34*, 577–590. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.