

# Assessment in Action: Investigating the Practices of Malaysian Physical and Health Education Teachers

Zuraimi Zakaria, Amirul Izzat Shazwan Ahmad Rizal,  
Ani Mazlina Dewi Mohamed

Faculty of Education, Universiti Teknologi MARA, Malaysia

Corresponding Author Email: zurai125@uitm.edu.my

To Link this Article: <http://dx.doi.org/10.6007/IJARPED/v13-i1/21046>

DOI:10.6007/IJARPED/v13-i1/21046

*Published Online:* 07 March 2024

## Abstract

Within the landscape of the physical and health education (PHE) curriculum in Malaysia, where instructional and assessment activities are framed by a school-based assessment framework, valid and reliable assessment practice is critical. However, there is a distinct lack of studies exploring Malaysian PHE teachers' assessment competencies and practice, resulting in considerable gaps in understanding how assessments are utilized to bolster the instructional process and the overall efficacy of the PHE curriculum implementation. Addressing these gaps, this study delves into teachers' assessment practices, encompassing traditional and alternative assessment methodologies, evidence-based instructional strategies, and the utilization of scoring tools. Employing a quantitative approach, a survey was administered to 63 PHE secondary school teachers across five urban districts of a state in Malaysia. Findings indicate a discrepancy between curriculum expectations and teachers' actual engagement with assessment practices, highlighting a lack of integration of assessment information to inform instructional strategies. Additionally, a clear preference for rubrics over other scoring methods was observed. Moreover, the study highlights missed opportunities in promoting critical thinking skills and utilizing informal assessment data to enhance ongoing instruction. These findings emphasize the need for comprehensive professional development and robust support systems to bolster effective assessment practices in Malaysian schools.

**Keywords:** Traditional Assessment, Alternative Assessment, Evidence-Informed Instruction, Scoring Tool, Testing

## Introduction

The assessment competencies of physical and health education (PHE) teachers stand as a critical determinant for the comprehensive development of students, encompassing cognitive, affective, psychomotor and social domains of learning. Beyond academic achievements, the goals of PHE extend to nurturing a healthy mind and lifestyle, ultimately contributing to students' overall quality of life and well-being (Keegan et al., 2019; Whitehead, 2019). As frontline educators tasked with this responsibility, PHE teachers play a critical role in designing and implementing assessments that accurately measure students' progress and

attainment of learning objectives. The validity and reliability of assessment practices are critical, ensuring that evaluations provide meaningful insights into students' abilities and areas for growth (Parsak & Sarac, 2021; Veloo, 2016; Veloo & Ali, 2016).

The literature measuring PHE teachers' assessment practices suggests that teachers' adeptness in assessment activities directly influences the quality of instruction Sitovskyi et al (2019); Van Munster et al (2019); Veloo & Krishnasamy (2017); student engagement Arefiev et al (2020); Parsak & Sarac (2021); Rafi & Pourdana (2023); and overall learning outcomes in PHE (Mohamed et al., 2019; Rongchan, 2020; Wee, 2019; Wee et al., 2021). Effective assessment strategies not only inform teachers about students' progress but also guide instructional decisions, allowing for tailored interventions to address individual learning needs (Arefiev et al., 2020; Colquitt et al., 2017). Furthermore, robust assessment practices promote a culture of continuous improvement, fostering students' self-awareness and motivation to excel (Rafi & Pourdana, 2023; Zakaria et al., 2023). However, shortcomings in teachers' assessment competencies can severely impede the efficacy of instruction, leading to inaccurate evaluations and missed opportunities for targeted interventions (Dudley et al., 2022).

The discourse surrounding the assessment practices of Malaysian PHE teachers highlights notable disparities, particularly when contrasted with the breadth of research and methodologies devoted to core academic subjects within the national curriculum (Wan Omar, 2019). Furthermore, while there exists a global interest in exploring the assessment competencies of PHE teachers, the corresponding body of literature at the local level appears comparatively scant (Atan et al., 2020; Wan Omar, 2019). This oversight is particularly significant given the backdrop of the Malaysian curriculum with slightly over ten years of its implementation. Studies assessing the effectiveness of this implementation have consistently highlighted low levels of execution (Abdullah et al., 2015; Lebar et al., 2013; Veloo & Krishnasamy, 2017), emphasizing the crucial role of teachers' assessment competency as a determining factor for effective teaching and learning activities (Mohamed et al., 2019; Veloo, 2016). This is especially pertinent considering the curriculum's foundation on a school-based assessment framework (Wee, 2019; Wee et al., 2021). Thus, the scarcity of studies into Malaysian PHE teachers' assessment practices emphasizes crucial gaps in understanding the holistic landscape of PHE education and its potential impact on student outcomes.

Aiming to address these gaps, this study aspired to examine PHE teachers' assessment practices in relation to the following research objectives:

- i. To investigate teachers' assessment practices in relation to processes in the implementation of traditional assessment;
- ii. To examine teachers' implementation of alternative assessment tasks;
- iii. To ascertain whether assessment data were used to inform instruction; and
- iv. To determine teachers' use of recording and scoring tools in their assessment processes.

The under-investigation of these areas contributes to a broader issue of insufficient evidence to guide policy formulation, curriculum design, and professional development in PHE, potentially stymieing efforts to elevate the status and effectiveness of PHE within the educational ecosystem. Addressing these research gaps is imperative for advancing a more in-depth understanding of PHE assessment practices, which is essential for promoting quality teaching, learning, and the realization of the full spectrum of educational outcomes that PHE aims to achieve.

## Literature Review

### *Physical and Health Education in the Malaysian Curriculum*

The Primary School Standard Curriculum and the Secondary School Standard Curriculum are key reforms within the Malaysian education system, initiated with the objective of enhancing educational standards and addressing the evolving demands of an increasingly dynamic global landscape. Introduced in phases beginning in 2011, the curriculum emphasizes a holistic, student-centred approach to learning, prioritizing the mastery of competencies alongside the cultivation of critical thinking and problem-solving abilities (Curriculum Development Division, 2017; Ministry of Education Malaysia, 2013). The curriculum highlights the Malaysian Ministry of Education's (MOE) dedication to nurturing balanced growth in students—intellectually, spiritually, emotionally, and physically—in accordance with the National Education Philosophy (Abu Hassan et al., 2020; Ministry of Education Malaysia, 2013).

This educational reform aligns with the aspirations of the Malaysia Education Blueprint 2013-2025 to produce graduates who are not only academically accomplished but also possess the skills necessary for employability. The integration of Higher Order Thinking Skills (HOTS) is largely a response to Malaysia's performance in international assessments such as the Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS), where Malaysian students have historically underperformed compared to their global counterparts (Ministry of Education Malaysia, 2013). By incorporating HOTS across various subjects and year levels, the objective was to transform students from passive recipients of knowledge into active learners who can apply knowledge in diverse contexts, critically analyse information, and devise innovative solutions to real-world challenges (Abu Hassan et al., 2020; Curriculum Development Division, 2017, 2018). Additionally, the focus on HOTS is considered essential in addressing both the immediate educational deficiencies identified through international assessments and the long-term economic goals of the nation (Ministry of Education Malaysia, 2013).

The Malaysian curriculum is underpinned by the School-based Assessment (SBA) framework, comprising four key components: Classroom-based Assessments (CBA), Psychometric Assessment, Physical Activity, Sports and Co-curriculum Assessment (PAJSK), and Centralized Assessment. CBA involves a range of instructional activities, with formative assessments and evidence-informed instruction serving as the guiding principles. PAJSK evaluates students' physical fitness, motor skills, and engagement in sports and co-curricular activities, emphasizing the importance of physical health and promoting a lifelong interest in physical activity. The Psychometric Assessment measures students' cognitive abilities and aptitudes, facilitating interventions that support intellectual growth. The Centralized Assessment consists of a standardized large-scale assessment administered by the Malaysian Examination Syndicate for form five students at the end of secondary school (Curriculum Development Division, 2017; Malaysian Inspectorate Group, 2019). The integration of these components within the SBA framework demonstrates Malaysia's dedication to a holistic educational approach that encompasses both academic and non-academic aspects of student development (Ministry of Education Malaysia, 2013).

The implementation of CBA is closely integrated with standard curriculum documents, which outline content standards and learning standards, thus providing a well-defined pathway for educational achievement through a learning progression known as Mastery Levels (ML). The ML is constructed based on the mastery of knowledge and skills outlined in the Instructional Objectives of Bloom's Taxonomy and is further delineated according to the mastery of each topic. This structure allows teachers to continuously scaffold student growth along the

progression and to report student achievement based on students' relative positions on the progression. Teachers utilize the ML to make informed judgments about students' learning abilities, guiding their decision-making process in tailoring instruction and intervention strategies to meet diverse learning needs (Curriculum Development Division, 2017, 2018). The application of ML in CBA facilitates a dynamic and responsive educational approach, wherein assessments are both informed by and integrated into daily instruction, fostering a culture of continuous learning and improvement (Ministry of Education Malaysia, 2013).

### ***Physical and Health Education (PHE) Teachers' Assessment Practices***

The landscape of research on Malaysian PHE teachers' assessment practices presents a stark contrast to the wealth of studies dedicated to other subject disciplines. Specifically, the body of literature concerning PHE teachers' proficiency in SBA, particularly CBA practices, remains markedly underdeveloped compared to the extensive exploration of examination-based core subjects (Wan Omar, 2019). Intriguingly, the existing investigations into the implementation of SBA by Malaysian PHE teachers predominantly delve into their perceptions (Sani et al., 2020) and the myriad challenges and constraints PHE teachers encounter in executing effective instruction (Kilue & Mohamad, 2017; Veloo, 2016; Wan Ismail & Muhamad, 2015; Wee & Chin, 2020; Yaakop et al., 2021). This highlights a significant research gap, with a noticeable dearth of studies examining the assessment practices employed by PHE teachers and a minimal focus on the intricacies of classroom assessment activities.

Within the limited availability of studies measuring Malaysian PHE teachers' assessment practices, those conducted report a low to moderate level of adoption. Lebar et al (2013) noted that PHE teachers lack proficiency in constructing test items and clarity regarding what constitutes effective assessments. Similarly, Veloo and Krishnasamy (2017) reported that teachers did not have a clear understanding of PHE assessment requirements, with this lack of understanding being particularly prevalent among non-option teachers. Abdullah et al (2015) revealed that only 30 percent of the PHE teachers in their study had the competency to develop assessment instruments. Additionally, Saad et al. (2016) observed a misalignment between PHE teachers' assessment practices and curriculum standards and expectations. They noted that assessments failed to comprehensively measure the intended learning standards and the practice lacked the test planning process. In a related study, Sani et al (2020) indicated that health education teachers predominantly fell within the range of being 'indecisive or agree' and 'sometimes to frequently' for items examining alignment, assessment components, administration, creativity, and innovation.

The efficacy of PHE teachers' assessment practice is influenced by a myriad of factors. These include student-related factors such as attitudes, motivation, and interest (Ali & Rauf, 2017; Ali et al., 2017; Wan Ismail & Muhamad 2015); resource-related factors like the scarcity of facilities and inadequate sports equipment (Mohamed et al., 2019; Wee, 2019; Wee et al., 2020); and school management factors involving the provision of support and effective governance (Gazali et al., 2022; Wee et al., 2021). Among these, teacher-related factors have been identified as the most significant (Ali & Rauf, 2017; Yaakop et al., 2021;). Central to effective instructional activities in PHE is teacher competency, which is associated with the level of knowledge and skills Mohamed et al (2019), experience Ali & Rauf (2017), creativity and innovativeness Wee et al (2021), as well as self-efficacy (Handrianto et al., 2024; Yaakop et al., 2021).

There is a pressing need for robust and comprehensive investigations into the assessment practices of Malaysian PHE teachers. A key metric in evaluating the effectiveness of PHE

curriculum implementation involves scrutinizing the discrepancies between current practices and established standards. Identifying these gaps is crucial to facilitate the provision of targeted training and professional development, as well as other concerted efforts aimed at enhancing the quality of PHE instruction and assessment. Additionally, understanding these gaps is instrumental in determining the effectiveness of the current curriculum's implementation, ensuring that it aligns with educational objectives and contributes to the holistic development of students.

### **Method**

The study employs a quantitative framework, with a survey serving as the primary instrument for data collection. A survey is notably efficacious for its ability to yield robust and reliable data across extensive populations. It offers a uniform and organized means of eliciting information from a wide array of respondents, thereby guaranteeing data collection homogeneity (Cresswell, 2013).

The instrument utilized in this study was adapted from the Assessment Practice Inventory by Zhang and Burry-Stock (2003), with the inclusion of additional items to more accurately reflect contemporary practices within the national curriculum. It consists of 36 items distributed over four sections. Excluding Section A, which collects demographic information of teachers through 6 items, the subsequent sections assess the frequency of teachers' engagement in assessment-related activities using a 5-point Likert scale. Data collection was conducted towards the end of the academic year, allowing teachers to report their involvement in each assessment activity over the span of a year. The instrument underwent a pilot test with six PHE teachers from two secondary schools, facilitating the refinement of items before the survey's administration. The instrument's overall reliability, as indicated by a Cronbach's Alpha score, was 0.923, demonstrating its high reliability for this research context.

63 secondary school Physical and Health Education (PHE) teachers from five urban districts in Selangor participated in the study. The quantitative data was analysed using descriptive statistics through the use of Statistical Package for the Social Sciences (SPSS) software (version 26).

### **Findings**

#### ***Demographic Information of Respondents***

The study garnered participation from 63 PHE teachers distributed across five urban districts in Selangor: Shah Alam (n=35, 56%), Subang Jaya (n=8, 13%), Petaling Jaya (n=11, 17%), Damansara (n=5, 8%), and Klang (n=4, 6%). Among these respondents, 40 (64%) were female, and 23 (36%) were male. The age profile indicated that 62 teachers were within the 31 to 60 age range, with a single respondent aged between 25 to 30 years. The predominant age group comprised teachers aged 36 to 45 years (n=32, 51%), followed by those aged 31 to 35 years (n=16, 25%), and 46 to 60 years (n=14, 22%). In terms of academic qualifications, 39 teachers (62%) held Bachelor's degrees, and 24 had achieved Master's degrees as their highest level of academic attainment; among them, 43 (68%) specialized in PHE, while 20 (32%) had other specializations. The duration of teaching experience among the participants varied, with a significant proportion (n=35, 56%) reporting 10 to 14 years of service. Furthermore, 10 teachers had 5 to 9 years of experience, 13 (21%) had been in the profession for 15 to 19 years, and a smaller group of 5 (8%) had exceeded 20 years of teaching experience.

***PHE Teachers' Traditional Assessment Practice***

The first research objective was to explore the practices of Physical and Health Education (PHE) teachers in conducting traditional assessments, specifically focusing on the procedural aspects of classroom testing. This investigation encompassed 13 items representing a comprehensive range of testing activities. The level of involvement in each activity was assessed using a 5-point Likert Scale, where teachers indicated the extent of their engagement as 'never,' 'occasionally,' 'for half of the assessments,' 'for almost all assessments,' or 'for each assessment' within a year. Table 1 presents the descriptive analysis for these items.

Table 1

***Descriptive Statistics for Teachers' Engagement in Traditional Assessment Activities***

How often have you engaged in the following activities this year?		mean	SD
1	Inform students in advance about assessments and their administration.	1.90	1.043
2	Inform students in advance about the test scoring process.	2.78	0.975
3	Teach students test taking skills.	2.35	1.358
4	Develop a table of specifications for test planning.	1.44	1.161
5	Ensure item types match the content assessed.	2.49	1.268
6	Include Lower Order Thinking Skills (LOTS) items in tests.	2.44	1.228
7	Include Higher Order Thinking Skills (HOTS) items in tests.	2.22	1.325
8	Map developed items to the complexity levels in the DSKP's Levels of Mastery ( <i>Tahap Penguasaan</i> ).	2.75	0.983
9	Prepare an answer scheme.	2.78	1.156
10	Conduct test following the school's and MOE's procedures.	2.83	1.185
11	Perform item analysis to evaluate test items' effectiveness.	2.33	1.380
12	Interpret item analysis indices.	2.37	1.360
13	Improve test items based on the item analysis.	2.49	1.306

The analysis of PHE teachers' engagement in traditional assessment activities offers valuable insights into their assessment practices. Adherence to the school's and MOE's procedures for conducting tests (mean: 2.83, SD: 1.186) emerged as the most commonly adopted traditional assessment activity. Furthermore, communicating with students about the test scoring process (mean: 2.78; SD: 0.975) and preparing an answer scheme (mean: 2.78; SD: 1.156) were among the most frequently implemented testing processes. The analysis suggests that while certain practices, such as procedural adherence and scoring transparency, were prioritized, the degree of uniformity in their application varied considerably. The higher mean value and lower standard deviation for informing students about the scoring process suggest that this practice was more deeply embedded in the assessment norms among PHE teachers, indicating that these areas were well-understood and consistently applied. Conversely, assessment activities related to test planning and communication with students about assessments and their administration were less commonly practiced, as indicated by lower mean values (1.44 and 1.90, respectively). The relatively low mean values signal a notable shortfall in the adoption of these assessment strategies. The moderate standard deviation values highlight variability in the execution of these practices, suggesting a lack of uniformity among teachers in their implementation.

The analysis reveals a moderate frequency in the application of specific elements of the testing process. Teaching students test-taking skills (mean = 2.35, SD = 1.358) and integrating HOTS items (mean = 2.22, SD = 1.325) both exhibited moderate mean values, indicating that these practices were adopted by some, but not all, teachers. The relatively high standard deviations suggest considerable variability in the consistency of these practices, highlighting a diversity in teaching strategies and potential differences in priorities or understandings of assessment objectives among teachers.

The analysis further revealed that teachers exhibited some consistency in incorporating (Lower Order Thinking Skills) LOTS items into their classroom tests, while the integration of HOTS items was less uniform. This variability indicates a potential gap in challenging students cognitively through assessments. Additionally, activities related to item analysis, interpretation, and improvement showed moderate levels of engagement, with mean values for these items ranging from 2.33 to 2.49. The relatively high standard deviations suggest a disparity in teacher practices, indicating that while some teachers were actively engaged in the test improvement process, others might not be. This points to a potential area for professional development to enhance assessment practices and ensure a more consistent application of higher-order cognitive challenges in assessments.

### ***PHE Teachers' Alternative Assessment Practice***

The second research objective aimed to determine the extent of PHE teachers' engagement in alternative assessment activities. This construct was measured using eight items on a 5-point frequency-oriented Likert Scale, with options ranging from 'never,' 'once in a while,' '3 to 4 times,' '5 times or more,' to 'for each assessment.' The results of the descriptive analysis are presented in Table 2.

Table 2

#### ***Descriptive Statistics for Teachers' Engagement in Alternative Assessment Activities***

How often have you engaged in the following activities this year?		mean	SD
1	Determine which learning domain (cognitive, affective, psychomotor) an assessment should address before selecting the specific assessment task.	2.65	1.050
2	Choose the right assessment task for the intended learning outcomes.	2.10	1.364
3	Select an assessment task that aligns with teaching styles.	2.62	1.170
4	Select an assessment task accommodating students' varied learning styles, preferences, and backgrounds.	1.71	1.197
5	Incorporate assessment task that targets learning in the affective domain.	1.95	1.464
6	Incorporate assessment task that targets skills acquisition in the psychomotor domain.	2.37	1.371
7	Use a readily available scoring tool without changes.	2.44	0.857
8	Provide students with a copy of scoring tools (e.g., rating scales, rubrics) before starting the assessment task.	2.62	1.084

The adoption of alternative assessment activities by PHE teachers demonstrated variability, reflecting the complexities of assessment and the diverse contexts in which they operate. The findings revealed a significant emphasis on aligning assessments with specific learning domains (mean: 2.65), highlighting the importance of ensuring assessments are meaningful

and accurately measure intended educational outcomes. Additionally, selecting assessment tasks that align with teaching styles was also frequently adopted (mean: 2.62), suggesting the importance of maintaining consistency between teaching and assessment methods. Furthermore, the findings showed a tendency among teachers to adopt practices that facilitated clarity and understanding of assessment expectations (mean: 2.62), likely due to the direct benefits of enhancing student engagement and performance. However, the standard deviation values for these items (1.050, 1.170, and 1.084 respectively) indicated moderate variability, suggesting that while there was general agreement on the importance of these practices, there were differences in the frequency of implementation.

Conversely, alternative assessment tasks that required more individualized adaptation and customization to student needs were the least frequently adopted (mean: 1.71; SD: 1.197). This low engagement might reflect the challenges faced by PHE teachers in tailoring assessments to the diverse needs of their students, possibly due to constraints in resources, time, or expertise in differentiated assessment strategies. The assessment within the affective domain (mean: 1.95; SD: 1.464) was also practiced less frequently compared to the assessment of skills pertaining to the psychomotor domain (mean: 2.37; SD: 1.371). The mean values indicate a moderate frequency of adopting these practices, with a slightly higher emphasis on psychomotor skills development. The relatively high standard deviations for both items, however, reveal significant variability in the frequency with which these assessment tasks were incorporated, suggesting that while some teachers regularly included these domains in their assessments, others may do so less frequently or not at all.

Alternative assessment activities involving the use of readily available scoring tools (mean: 2.44; SD: 0.857), incorporating assessment tasks targeting psychomotor-focused learning (mean: 2.37; SD: 1.371), and selecting suitable assessment tasks for intended learning outcomes (mean: 2.10; SD: 1.364) were moderately adopted. The variability, as indicated by the standard deviations, suggested a significant range in the consistency of these practices among teachers.

### ***PHE Teachers' Evidence-informed Instruction Practice***

The third research objective explored the degree to which PHE teachers utilized evidence of learning to inform instruction and intervention strategies. This construct was primarily based on items that represented the incorporation of evidence of learning from activities focusing on informal assessments. The emphasis on informal assessments allowed for the measurement of evidence-based practice, as teachers continuously integrate immediate evidence from informal assessment activities to modify and enhance the teaching and learning process. Teacher responses were recorded based on a 5-point Likert Scale ranging from 'never,' 'once in a while,' '4-5 times a year,' '2-3 times a month,' and 'each assessment.' The results of the descriptive analysis are presented in Table 3.



Table 3

*Descriptive Statistics for Teachers' Use of Assessment Data*

How often have you engaged in the following activities this year?		mean	SD
1	Utilize quiz responses to evaluate and improve teaching effectiveness.	2.02	1.314
2	Provide feedback to students based on evidence from question-and-answer session.	2.40	1.326
3	Provide feedback to students based on evidence from teacher observation during instructional activities.	2.35	1.310
4	Plan instruction based on the Mastery Level (TP) identified from assessments.	2.29	1.337
5	Determine students' level of learning based on combined evidence from assessment data and multiple other sources (e.g. attendance records, socioeconomic background, co-curricular data, etc.).	2.29	1.313

The analysis suggests that the instructional practices of PHE teachers were not predominantly evidence-based. The relatively low mean values across all items indicate missed opportunities in leveraging assessment data to enhance instruction. The consistency of these practices varied significantly among teachers, as evidenced by the high standard deviation values, implying that only a subset of teachers might be regularly engaged in these practices compared to others. This variation highlights the need for targeted professional development to promote the integration of assessment data into instructional strategies.

Within the spectrum of infrequently adopted evidence-informed practices, the provision of feedback based on data from question-and-answer sessions was somewhat more frequently used (mean: 2.40, SD: 1.326). In contrast, the utilization of quiz responses to inform teaching effectiveness was the least frequent practice (mean: 2.02; SD: 1.314). Furthermore, teacher practice was not data-driven, as multiple sources of data were not utilized to inform and enhance instruction (mean: 2.29; SD: 1.313).

***PHE Teachers' Use of Scoring Tools***

The fourth research objective focused on examining the adoption of recording and scoring tools by PHE teachers. Four items representing specific tools were presented on a five-point Likert Scale, with options ranging from 'never,' 'once in a while,' '3-4 times,' '5 times or more,' and 'for each assessment.' Table 4 presents the descriptive analysis of these items.

Table 4

*Descriptive Statistics for Teachers' Adoption of Scoring Tools*

How often have you used the following recording/scoring tools this year?		mean	SD
1	Log book	1.52	0.759
2	Checklist	1.73	0.937
3	Rating scale	1.75	0.967
4	Rubric	2.21	0.864

The analysis of PHE teachers' utilization of scoring tools revealed distinct patterns in frequency and consistency of use across four tools. The scoring rubric recorded the highest mean value (mean: 2.21; SD: 0.864), suggesting it was the most frequently used tool among teachers, albeit with moderate variability in their usage. The higher mean indicated a preference or greater reliance on rubrics for assessment, possibly due to their comprehensive

nature in evaluating student performance. The moderate standard deviation reflected some consistency in use, yet indicated that the extent of rubric usage varied among teachers.

The log book recorded the lowest mean score (mean: 1.52; SD: 0.759), indicating that it was the least frequently used tool, with relatively low variability in their usage. The low mean suggested a limited reliance on log books for recording and scoring, possibly due to their perceived inefficiency or the availability of more effective tools. The lower standard deviation indicated a higher level of consistency among teachers in their infrequent use of log books, suggesting a general agreement on their limited utility in the assessment process.

Checklists and rating scales recorded similar mean values (1.73 and 1.75, respectively) with slightly higher standard deviations (0.937 and 0.967, respectively) compared to the log book. These tools were used somewhat frequently but with greater variability in usage compared to rubrics. The close mean values suggested that both tools were considered useful, yet the higher standard deviations indicated significant inconsistency in how frequently teachers utilized them. This variability could reflect personal preference, student needs, or the specific context of the assessment, highlighting the diverse approaches to assessment adopted by PHE teachers.

### **Discussion**

The study delved into the assessment practices of physical and health education (PHE) teachers, examining their engagement with traditional assessment, alternative assessment, evidence-informed instruction, and the use of recording and scoring tools. The findings indicated that although teachers moderately engaged in activities related to the testing process and alternative assessment tasks, their instructional practices were not evidence-based. The data from assessments were not effectively utilized to enhance instruction. While rubrics were more commonly employed in the scoring process, the use of other scoring tools was infrequent, indicating a lack of diversity in recording and evaluating evidence of learning. The findings revealed a varied engagement in traditional assessment practices among PHE teachers. High adherence to procedural norms, such as following school and MOE guidelines and maintaining transparency in scoring processes, highlighted a commitment to reliability and fairness in assessments. These practices, characterized by higher mean values and lower standard deviations, suggested a relatively uniform application and possibly reflected an institutional norm that valued structured assessment frameworks. In addition, the effort to map developed items to the complexity levels in the curriculum's Levels of Mastery indicated an intention to align assessments with standardized educational goals and benchmarks. This practice ensured that assessments were not arbitrary but were designed to measure students' progress against recognized standards of achievement.

However, the lower engagement in test planning indicated potential areas of weakness. Completing a table of specifications is a form of test planning process and is a critical aspect of traditional assessment. Its practice ensures content validity, one of the precursors to high-quality assessment (Barnes et al., 2014; Erlinawati & Muslimah, 2021). Additionally, communicating about assessments and their administration to students demonstrates teachers' commitment to assessment transparency (Herman & Cook, 2019; Tohfighi & Safa, 2023). The underutilization of these practices could lead to missed opportunities for aligning assessments with learning objectives and fostering a more inclusive and transparent educational environment (Dudley et al., 2022; Tohfighi & Safa, 2023). Moreover, the relatively low engagement in improving test items based on item analysis highlighted a critical gap in using data to enhance assessment quality (Zakaria & Abdul Latif, 2023). This suggested a

missed opportunity for iterative improvement of assessments, which is essential for ensuring that they remain relevant, fair, and effectively measure what they are intended to (Barnes et al., 2021; Erlinawati & Muslimah, 2021; Zakaria & Abdul Latif, 2023).

The inclusion of both LOTS and HOTS items suggested an awareness among some teachers of the need to assess a range of cognitive skills, from basic recall to critical thinking and problem-solving. This was a positive step towards developing assessments that fostered a deeper understanding of the subject matter and encouraged students to apply their knowledge in various contexts. However, assessments provided more opportunities for students to engage with LOTS with minimal exposure to HOTS. The moderate mean values and relatively high standard deviations for these items suggested significant variability among teachers in how these practices were applied.

Contrary to studies that identified a low level of assessment implementation (Saad et al., 2016; Abdullah et al., 2015; Lebar et al., 2013), the findings of this study align with those of Sani et al (2020), which reported a moderate adoption of the measured assessment practices by teachers. However, similar to the findings of Saad et al (2016), the PHE teachers in this study largely did not engage in the test planning process. This gap in test planning and item refinement observed in the study also reflects recent critiques that, despite awareness, low competency level and practical challenges hinder optimal assessment design and improvement (Heitink et al., 2019; Mohsin, 2022).

The findings of this study regarding alternative assessment practices indicate an awareness of the need to tailor assessments to specific learning domains and teaching styles, though with varying degrees of application. This variability, along with the less frequent use of individualized assessment tasks, suggests potential challenges teachers face in adapting assessments to meet the diverse needs of students. The focus on assessing psychomotor skills aligns well with the physical education aspect of PHE, highlighting the recognition of the importance of skill acquisition and physical competency. Nevertheless, the less frequent incorporation of affective domain assessments suggested a potential undervaluing of health education's behavioural, emotional, and social learning outcomes. This uneven emphasis can lead to a less holistic development of students, overlooking the importance of attitudes, values, and self-awareness in health education component. It also pointed to a potential gap in addressing the full spectrum of educational objectives within PHE.

The variability and moderate engagement in alternative assessment tasks among PHE teachers reflect a broader educational trend towards diversified assessment strategies that encompass a wide range of skills and learning domains. Studies by Mohamed et al (2019); Veloo (2016); O'Neill and Padden (2021) highlight the increasing recognition of the importance of such practices to capture the full range of student learning, advocating for more consistent application and teacher training in alternative assessment methods. The emphasis on psychomotor skills assessment in PHE, while less so on affective domain assessments, mirrors concerns raised in the literature about the need for a more holistic approach to evaluating student outcomes (Fernández-Río et al., 2017).

The lower mean values across evidence-informed instruction practices highlighted a significant gap in using assessment data to inform and enhance teaching strategies. This deficiency in evidence-based practice not only limited the potential for targeted instructional interventions but also reflected a broader issue within the educational assessment culture that undervalued the dynamic use of data for continuous improvement. The variability in these practices suggested a disparity in teacher capabilities or motivations to integrate assessment data into instructional decision-making, highlighting a critical area for

professional development. The gap in utilizing assessment data to inform instruction, as noted in the study, is a significant issue that recent literature continues to address. Mandinach and Schildkamp (2021); Van Geel et al (2017); Zakaria et al (2023) are among the authors who emphasize the critical role of data literacy for teachers in leveraging assessment information to enhance teaching practices. The variability in teachers' abilities to integrate assessment data into their instruction emphasizes a persistent challenge in education: fostering a culture of evidence-informed practice (Zakaria & Abdul Latif, 2023; Zakaria et al., 2023).

The utilization of scoring tools by physical and health education (PHE) teachers demonstrates a noticeable preference for rubrics over other assessment tools such as logbooks, checklists, and rating scales. This preference suggests an acknowledgment of the comprehensive and reliable judgment of student abilities that rubrics provide. However, relying solely on one type of assessment tool can lead to reliability issues, as different learning outcomes may require varied forms of evaluation to consistently measure student progress, growth, and achievement (Rongchan, 2020). The underutilization of alternative scoring tools indicates a potential gap in capturing the full spectrum of student learning. For instance, in assessing physical skills where quick and practical scoring tools are needed, rubrics may not be the most appropriate choice. The recording and scoring process of physical and fitness tests can be more reliably accomplished through the use of checklists and rating scales (Wood & Pugh, 2019). Similarly, logbooks could enable teachers to track students' growth over time, allowing for more accurate judgments in teachers' decisions regarding learning ability (NarjesMoasheri et al., 2022; Nugraheni et al., 2023). While the preference for rubrics is beneficial, it may indicate a need for broader training and support to encourage a more diversified use of assessment tools. This would enhance the ability to capture a wide range of student achievements and learning outcomes (Panadero & Jonsson, 2013; Rongchan, 2020).

The findings also revealed a lack of personalization and customization in both traditional and alternative assessment practices. The reliance on only one type of scoring tool suggested a homogeneous approach to scoring. This approach not only limited the ability of assessments to accurately measure individual learning progress in a manner that reflected students' unique contexts and learning preferences but also hindered students' engagement with the assessments in meaningful ways (Arefiev et al., 2020; Rafi & Pourdana, 2023). High-quality assessments, apart from having valid and reliable assessment tasks, procedures, administration, and scoring processes, should also offer items and tasks spanning an array of depths and complexities (Sitovskyi et al., 2019; Van Munster et al., 2019). This ensures that all students have the opportunity to demonstrate their mastery in ways that align with their individual strengths and areas for growth (Arefiev et al., 2020; Colquitt et al., 2017). A differentiated assessment approach not only acknowledges but also values the diverse learning styles and paces of students, promoting a more equitable and comprehensive assessment strategy (Sitovskyi et al., 2019; Van Munster et al., 2019). Consequently, the absence of personalized and customized assessment items and tasks represented a missed opportunity to fully engage and assess each student's unique learning journey, potentially hindering the overall educational outcomes (Colquitt et al., 2017; Dudley et al., 2022).

### **Limitations**

This study is subject to several limitations. The primary limitation arises from the exclusive reliance on surveys for data collection. While surveys are adept at collecting quantitative data efficiently, their capability to deeply explore the subtleties and intricacies inherent in PHE teachers' assessment practices is somewhat restricted. A more robust understanding would

be attainable through a mixed-methods research approach, blending both qualitative and quantitative methodologies. The inclusion of qualitative data could either support or contest the quantitative findings, yielding a more comprehensive perspective of teachers' practices. Additionally, incorporating qualitative insights would enable investigation into the impacts of assessment practices on teaching effectiveness and student learning outcomes. Moreover, the employment of diverse data collection techniques, including but not limited to interviews, observational methods, and document analysis, would facilitate methodological triangulation, thereby significantly augmenting the validity and reliability of the research outcomes. Lastly, the sample is limited to secondary school PHE teachers from urban districts of Selangor, which may not fully represent the diversity of teaching contexts across different regions or educational levels. This geographical and contextual limitation restricts the generalizability of the findings to broader educational settings.

### **Implications, Contributions and Future Research**

The findings critically emphasize the need for a more balanced and comprehensive approach to assessment, stressing the importance of aligning assessments with diverse learning domains and teaching styles. To improve the quality of assessments and enhance the teaching and learning process, several specific recommendations can be made. Firstly, professional development opportunities should be provided to PHE teachers to enhance their assessment literacy, focusing on utilizing data for instructional improvement and diversifying assessment strategies. Secondly, policymakers and curriculum developers should prioritize the provision of support and resources for teacher training, with a focus on developing teacher competencies in item construction and assessment methodologies. Additionally, there is a need for promoting evidence-informed instruction tailored to student needs, ensuring that teaching strategies align with assessment practices. Moreover, encouraging collaborative approaches among PHE teachers to share best practices and experiences can foster a culture of continuous improvement in assessment practices. By implementing these recommendations, the quality and effectiveness of assessments in PHE can be significantly enhanced, ultimately leading to improved student learning outcomes and holistic development.

Building on these findings, future research should focus on expanding the understanding of assessment practices in physical and health education (PHE) by exploring the integration of evidence-informed instructional strategies and the impact of diverse assessment tools on student learning outcomes. Qualitative studies are recommended to delve deeper into the reasons behind PHE teachers' preferences for certain assessment tools and their perceptions of the challenges in implementing a broader range of assessment strategies. Additionally, comparative studies across different educational contexts and cultures could offer a global perspective on best practices in PHE assessment. Lastly, exploring the use of technology in assessment practices could uncover innovative approaches to evaluating student performance in PHE, potentially transforming traditional assessment paradigms to better accommodate the complexities of learning in this domain.

### **Conclusion**

The overall findings reveal a complex landscape of assessment practices among PHE teachers, characterized by strengths in procedural adherence and scoring transparency but weakened by gaps in test planning, communication, and evidence-informed instruction. The variability in the adoption and application of both traditional and alternative assessment practices

suggests a need for more consistent and comprehensive training and support for teachers. Strengthening these areas could significantly enhance the alignment between assessment practices and educational objectives, ultimately fostering a more inclusive, effective, and reflective learning environment. Future studies should explore strategies to address these gaps, focusing on professional development, the integration of technology in assessment, and the cultivation of a culture that values continuous improvement and evidence-informed practices.

The study's theoretical contribution lies in its exploration of the integration of traditional and alternative assessment methodologies, offering insights into the challenges and preferences in assessment practices. Contextually, the findings highlight the need for comprehensive professional development to enhance teachers' assessment competencies, particularly in utilizing assessment data to inform instruction and promoting a more diversified approach to assessment. The study significantly contributes to the existing knowledge by addressing the underexplored area of PHE assessment practices in Malaysia, shedding light on the complexities of implementing a school-based assessment framework and its impact on instructional effectiveness and student outcomes.

### Acknowledgement

This paper is part of a research project funded by the Universiti Teknologi MARA Selangor Branch Research Fund under the DUCS-FAKULTI scheme [grant number: 600-UiTMSEL(PI.5/4)(115/2022)]

### References

- Abu Hassan, F. N., Mohamad, Q. A., Rosli, M. N. A., Ajmain@Jima'ain, M. T., & Azuddin, S. K. Y. (2020). The implementation of higher order thinking skills (HOTS) in Malaysia secondary school: Post PISA 2009. *International Journal of Psychosocial Rehabilitation*, 24(5), 5510-5517.
- Abdullah, N., Idris, N., Hamzah, M. S. G., & Sembak, S. (2015). Planning and implementation of School-Based Assessment (SBA) among teachers. *Social and Behavioral Sciences*, 211, 247-254. <https://doi.org/10.1016/j.sbspro.2015.11.031>
- Ali, S. K. S., & Rauf, P. A. (2017). Tahap perancangan bahan sumber, kemudahan dan peralatan pengajaran dalam kalangan guru pendidikan jasmani. *International Journal of Education, Psychology and Counselling*, 2(5), 170-176.
- Ali, S. K. S., Rauf, P. A., & Salimin, N. (2017). Hubungan antara pengalaman mengajar dan perancangan pengajaran dalam kalangan guru pendidikan jasmani tingkatan empat. *International Journal of Education Psychology and Counselling*, 2(6), 268-277.
- Arefiev, V., Tymoshenko, O., Malechko, T., Domina, Z., Bezkopyiny, O., Dutchak, Y., Riabchenko, V., Garmata, O., Griban, G., Rusanivskiy, S., Melnychuk, V., Bloshchynsky, T., & Prontenko, K. (2020). Methodology of differentiation of health-improving classes in physical education for primary school students. *International Journal of Applied Exercise Physiology*, 9(7), 134-143.
- Atan, N. A., Mohamad, Z., Yaacob, N., & Shafie, N. (2020). Amalan pentaksiran bilik darjah (PBD) bagi pentaksiran autentik dalam kalangan guru sekolah rendah. *Proceedings of International Conference on The Future of Education (IConFed) 2020*, 245-253. [https://www.ipgktb.edu.my/iconfed/compilation\\_of\\_paper\\_iconfed2018\\_volum2/index.html](https://www.ipgktb.edu.my/iconfed/compilation_of_paper_iconfed2018_volum2/index.html)

- Barnes, N., Fives, H., & Krause, E. S. (2014). Using a table of specifications to improve teacher-constructed traditional tests: An experimental design. *Assessment in Education: Principles, Policy & Practice*, 21(1), 90-108.  
<https://doi.org/10.1080/0969594X.2013.808173>
- Colquitt, G., Prtchard, T., Johnson, C., & McCollum, S. (2017). Differentiating instruction in physical education: Personalization of learning. *Journal of Physical Education, Recreation and Dance*, 88(7), 44-50.  
<http://dx.doi.org/10.1080/07303084.2017.1340205>
- Cresswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks.
- Curriculum Development Division. (2017). *Report of the implementation of Primary School Standard Curriculum (revised 2017), and Secondary School Standard Curriculum (revised 2017)*. Ministry of Education Malaysia.
- Curriculum Development Division. (2018). *Standard Curriculum and Assessment Document: Physical Education Health Education*. Ministry of Education Malaysia.
- Dudley, D., Mackenzie, E., Van Bergen, P., Cairney, J., & Barnett, L. (2022). What drives quality physical education? A systematic review and meta-analysis of learning and development effects from physical education-based interventions. *Frontiers*, 13, <https://doi.org/10.3389/fpsyg.2022.799330>
- Erlinawati, E., & Muslimah, M. (2021). Test validity and reliability in learning evaluation. *Bulletin of Community Engagement*, 1(1), 26-31.
- Fernandez-Rio, J., Sanz, N., Fernandez-Cando, J., & Santos, L. (2017) Impact of a sustained cooperative learning intervention on student motivation. *Physical Education and Sport Pedagogy*, 22(1), 89-105, <https://doi.org/10.1080/17408989.2015.1123238>
- Gazali, N., Saad, N., Cendra, R., Ali, S. K. S., & Ripa'i, R. (2022). Leadership in physical education: Systematic review of the last five years. *Jurnal SPORTIF Jurnal Penelitian Pembelajaran*, 8(4), 367-386. [http://dx.doi.org/10.29407/js\\_unpgri.v8i4.18226](http://dx.doi.org/10.29407/js_unpgri.v8i4.18226)
- Handrianto, C., Jusoh, A. J., Abdul Rashid, N., Suzaily, W., Abdullah, A., Khatim, M. H., & Rahman, M. A. (2024). Teacher's self-efficacy (TSE) and teaching competency (TC) of Malaysian secondary school teachers in drug education. *International Journal of Instruction*, 17(2), 219-236.
- Heitink, M., Van der Kleij, F., Veldkamp, B., Schildkamp, K., & Kippers, W. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review*, 17, 50–62. <https://doi.org/10.1016/j.edurev.2015.12.002>
- Kilue, D., & Mohamad, T. A. (2017). Cabaran pengajaran subjek pendidikan jasmani di sekolah menengah di Malaysia. *Journal of Nusantara Studies*, 2(2), 53-65.  
<https://doi.org/10.24200/jonus.vol2iss2pp53-65>
- Lebar, O., Mokshein, S. E., Daud, A. F., Rahmat, A., & Baharin, M. O. H. (2013). *Pembinaan standard Pentaksiran Berasaskan Sekolah bagi sekolah rendah di Malaysia*. Tanjung Malim, Malaysia.
- Malaysian Inspectorate Group. (2019). *Physical education health education teaching guide book*. Ministry of Education Malaysia.
- Mohamed, A. M. D., Amri, S., Kok, L. Y., & Abdullah, B. (2019). Factors influencing the implementation level of physical education in Selangor primary schools. *International Journal of Academic Research in Business and Social Sciences*, 9(13), 471-479.  
<http://dx.doi.org/10.6007/IJARBS/v9-i13/6898>

- Mohsin, M. (2022). Implementation of physical education curriculum by form one teachers in schools. *Proceedings International Conference on Education 2022*. [https://www.researchgate.net/publication/363470614\\_Implementation\\_of\\_Physical\\_Education\\_Curriculum\\_by\\_Form\\_One\\_Teachers\\_in\\_Schools](https://www.researchgate.net/publication/363470614_Implementation_of_Physical_Education_Curriculum_by_Form_One_Teachers_in_Schools)
- Mandinach, E. B., & Schildkamp, K. (2021). Misconceptions about data-based decision making in education: An exploration of the literature. *Studies in Educational Evaluation*, 69. <https://doi.org/10.1016/j.stueduc.2020.100842>
- Ministry of Education Malaysia. (2013). *Malaysia Education Blueprint 2013-2025 (Preschool to post-secondary education)*. Ministry of Education Malaysia. <http://eprints.iab.edu.my/v2/965/1/malaysia%20education%20blueprint%202013-2025.pdf>
- Moasheri, N. B., Miri, M. R., & Sharifzadeh, G. (2022). Comparison of instructors' and students' satisfaction with traditional and logbook evaluation methods in internship of public health. *Health Technology Assessment in Action*, 6(1). <https://doi.org/10.18502/htaa.v6i1.11129>
- Nugraheni, H., Murwani, R., Shaluhayah, Z., & Widjanarko, B. (2023). Logbook as parental mentoring media in daily physical activity of elementary school student. *Journal of Survey in Fisheries Sciences*, 1, 1-13. <https://sifisheressciences.com/journal/index.php/journal/article/view/723/705>
- O'Neill, G., & Padden, L. (2022) Diversifying assessment methods: Barriers, benefits and enablers, *Innovations in Education and Teaching International*, 59(4), 398-409, <https://doi.org/10.1080/14703297.2021.1880462>
- Panadero, E., & Johnsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144. <https://doi.org/10.1016/j.edurev.2013.01.002>
- Parsak, B., & Sarac, L. (2021). Health-related physical activity report card applications: an initial examination of physical education teachers' views. *Pamukkale Journal of Sport Sciences*, 12(3), 21-40. <https://doi.org/10.54141/psbd.982686>
- Rafi, F., & Pourdana, N. (2023). E-diagnostic assessment of collaborative and individual oral tiered task performance in differentiated second language instruction framework. *Language Testing in Asia*, 13(6). <https://doi.org/10.1186/s40468-023-00223-7>
- Rongchan, L. (2020). *Rubrics for scoring, interpretations and decision-making. Assessment rubrics decoded*. Routledge. <http://dx.doi.org/10.4324/9780429022081-5>
- Sani, F. M. A., Rahmat, A., Elumalai, G., & Nadzalan, M. A. (2020). Teachers' perspectives and practices of assessment in health education in primary school. *Indian Journal of Public Health Research and Development*, 11(6), 1243-1246. <http://dx.doi.org/10.37506/ijphrd.v11i6.9971>
- Saad, L., Jani, J., Rahmat, A., & Norkhalid, S. (2016). Proses reka bentuk model pentaksiran pembelajaran bagi permainan kategori serangan dalam pendidikan jasmani tahun empat. *Prosiding Konferens Kebangsaan Pendidikan Abad Ke-21*. Pejabat Pendidikan Daerah Larut Matang dan Selama.
- Schildkamp, K., Smit, M., & Blossing, U. (2019). Professional development in the use of data: From data to knowledge in data teams. *Scandinavian Journal of Educational Research*, 63(3), 393–411. <https://doi.org/10.1080/00313831.2017.137635>
- Sitovskyi, A., Maksymchuk, B., Kuzmenko, V., Nosko, Y., Korytko, Z., Bahinska, O., Marchenko, O., Matviichuk, T., Solovyov, V., Khurtenko, O., Zhorova, I., & Maksymchuk, I. (2019).



- Differentiated approach to physical education of adolescents with different speed of biological development. *Journal of Physical Education and Sport*, 19(3), 1532-1543.
- Tohfighi, S., & Safa, M. A. (2023). Fairness in classroom language assessment from EFL teachers' perspective. *Teaching English as a Second Language Quarterly*, 42(2), 81-110. <https://doi.org/10.22099/tesl.2023.46825.3173>
- Van Geel, M., Keuning, T., Visscher, A., & Fox, J. (2017). Changes in educators' data literacy during a data-based decision-making intervention. *Teaching and Teacher Education*, 64, 187-198. <https://doi.org/10.1016/j.tate.2017.02.015>
- Van Munster, M., Lieberman, L. J., & Grenier, M. (2019). Universal design for learning and differentiated instruction in physical education. *Adapted Physical Activity Quarterly*, 36(3), 1-19. <http://dx.doi.org/10.1123/apaq.2018-0145>
- Veloo, A. V. (2016). Physical education teachers' challenges in implementing School based Assessment. *International Review of Management and Marketing*, 6(8), 48-53.
- Veloo, A., & Krishnasamy, H. (2017). School-based Assessment in the context of secondary school physical education teachers in Malaysia. *PEOPLE: International Journal of Social Sciences*, 3(2), 2122-2134. <https://dx.doi.org/10.20319/pijss.2017.32.21222134>
- Veloo, A., & Ali, M. R. (2016). Teachers' challenges in the implementation of school-based assessment in physical education. *The European Proceedings of Social and Behavioural Sciences*, 6(58), 737-743. <https://doi.org/10.15405/epsbs.2016.08.104>
- Wan Ismail, W. A., & Muhamad, T. A. (2015). Hubungan faktor penghalang terhadap penglibatan pelajar dalam Pendidikan jasmani di sekolah menengah. *Proceedings Interantional Conference on Language, Education and Innovation 2015*, 22-29. <https://www.icsai.org/procarch/1iclei/1iclei-49.pdf>
- Wan Omar, S. H. (2019). Pengetahuan, kemahiran, sikap dan masalah guru dalam melaksanakan pentaksiran bilik darjah Bahasa Melayu di sekolah rendah. *Jurnal Pendidikan Bahasa Melayu*, 9, 56-67.
- Wee, E. H. (2019). Implementation of physical education programme in Malaysian primary schools. *Pertanika Journal Social Sciences and Humanities*, 27(53), 1-15.
- Wee, E. H., Cheng, W. F., & Chin, N. S. (2021). Teachers' perceived barriers to implementation of physical education: Examining the administration of physical education programme and the provision of non-human resources. *Collegium Antropologicum*, 45(3), 191-200.
- Wee, E. H., & Chin, N. S. (2020). Perceived barriers to the provision of physical education in Malaysian primary schools. *IOSR Journal of Sports and Physical Education*, 7(4), 22-28. <http://dx.doi.org/10.9790/6737-07042228>
- Whitehead, M. (2019). *Definition of physical literacy. Physical literacy across the world*. Routledge. <https://doi.org/10.4324/9780203702697-2>
- Wood, T. J., & Pugh, D. (2019). Are rating scales really better than checklists for measuring increasing levels of expertise? *Medical Teacher*, 42(4), 1-6. <http://dx.doi.org/10.1080/0142159X.2019.1652260>
- Yaakop, N., Lian, D. K. C., & Yasin, R. (2021). Challenges in teaching of physical education subject in Malaysian primary schools. *Prosiding Seminar Kebangsaan Pendidikan Negara (SKEPEN) 2021*, (pp. 434-448). Universiti Kebangsaan Malaysia.
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom Assessment Practices and Teachers' Self-Perceived Assessment Skills. *Applied Measurement in Education*, 16(4), 323-342. [https://doi.org/10.1207/S15324818AME1604\\_4](https://doi.org/10.1207/S15324818AME1604_4)
- Zakaria, Z., Ab. Wahid, N. T., & Abdul Latif, A. (2023). Data literacy competencies for informed classroom assessment practice: challenges and measures. *International Journal of*

*Academic Research in Progressive Education and Development*, 12(3), 2121-2136.  
<http://dx.doi.org/10.6007/IJARPED/v12-i3/19320>

Zakaria, Z., & Abdul Latif, A. (2023). Obstacles to effective data-based interventions: A systematic review of teachers' data literacy. *International Journal of Academic Research in Progressive Education and Development*, 12(3), 1708-1726.  
<http://dx.doi.org/10.6007/IJARPED/v12-i3/19221>