

Assessment Under Duress: Challenges to Physical Education Assessments and the Emergence of Compromised Norms

Zuraimi Zakaria¹, Ani Mazlina Dewi Mohamed¹, Diyana Zulaika Abdul Ghani², Mohammad Arif Ali³

¹Faculty of Education, Universiti Teknologi MARA, Malaysia, ²Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Malaysia, ³Faculty of Sports Science, Universitas Negeri Semarang, Indonesia

Corresponding Author Email: zurai125@uitm.edu.my

To Link this Article: <http://dx.doi.org/10.6007/IJARPED/v13-i3/21675>

DOI:10.6007/IJARPED/v13-i3/21675

Published Online: 16 June 2024

Abstract

This study investigated the challenges faced by Physical and Health Education teachers in implementing Classroom-based Assessment (CBA) practices, addressing a research gap in localized and practical constraints in slightly underrepresented regions of Malaysia. It sought to answer two research questions: what issues obstructed effective PE assessments, and how did these persistent challenges shape teachers' assessment norms and practices? Utilizing an explanatory qualitative research design, the study involved in-depth interviews with four teachers from Sabah and Sarawak. The thematic analysis revealed five main challenges: the complexity of CBA requirements, inherent complexities of PE assessments, lack of standardized assessment approaches, resource constraints, and administrative issues. Key findings indicated that these challenges led to compromised assessment practices, such as grade inflation and score averaging. The study highlighted the need for comprehensive professional development programs to enhance teachers' assessment knowledge and skills, fostering collaboration and consistent assessment practices. Recommendations included peer teaching, mentoring programs, and better resource allocation to improve PE assessments' effectiveness and alignment with educational objectives.

Keywords: Classroom-based Assessment, Physical Education, Issues and Challenges, Resource Constraint, Assessment Approaches

Introduction

Implementing valid and reliable assessments in Physical and Health Education (PHE) presents significant challenges due to the complexity of the discipline and the demanding skill sets required. Effective assessment critically relies on teachers' mastery of assessment knowledge and skills DeLuca et al (2019); Zakaria et al (2024), coupled with their pedagogical content knowledge (PCK) (Veloo & Ali, 2016; Wee et al., 2020). This expertise is essential for accurately measuring both tangible physical skills and salient aspects such as student engagement and

behavioural changes towards health (Kilue & Mohamad, 2017). However, studies measuring PHE teachers' assessment knowledge and skills revealed low level of mastery (Veloo & Krishnasamy, 2017; Zakaria et al., 2024) with the reliance on traditional assessment tools and strategies (Mohamed et al., 2019; Veloo, 2016; Zakaria et al., 2024). Often, teachers struggle with inadequate training in specialized assessment strategies (Veloo & Krishnasamy, 2017), leading to assessment norms that are not aligned with the practice delineated in the curriculum (Cui & Lei, 2015; Zhao et al., 2023). The subjective nature of many physical education assessments can also contribute to inconsistencies if not underpinned by robust pedagogical understanding (Veloo & Ali, 2016; Zakaria et al., 2024; Zhao et al., 2023).

This paper explores the challenges faced by PHE teachers in Malaysia as they implemented Classroom-based Assessment (CBA), a critical element of the Malaysian School-based Assessment (SBA) system that focuses on teaching and learning activities. It specifically examines the factors that impeded the execution of valid and reliable assessments and investigates how prolonged exposure to these issues led to the development of compromised norms in teachers' assessment practices. Utilizing an explanatory qualitative research design, the study involved four teachers from the western Malaysian states of Sabah and Sarawak, with data collected through in-depth interviews. The findings presented in this paper address two research questions:

- i. What were the issues and challenges that obstructed the effective implementation of valid and reliable PE assessments?
- ii. How did persistent and unaddressed challenges influence the shaping of teachers' assessment norms and practices?

The proposed study, which involves in-depth interviews with four teachers from Sabah and Sarawak, addresses a crucial research gap by offering a detailed qualitative exploration of assessment challenges in these underrepresented regions. Despite its relatively small sample size, the study promises valuable insights into the localized and practical constraints that shape assessment norms, a dimension not typically explored in depth by other studies.

The study is significant as it provides insights into the potential development of more effective, context-sensitive interventions, and enhances understanding of teacher assessment behaviours, particularly in these specific geographic areas. Identifying the barriers that teachers experience and understanding their influence on assessment practices is the first step towards the development of effective, targeted intervention strategies. It also helps to address the gaps in effective instructional and assessment practices as measures to ensure that CBA is implemented in accordance with the curriculum aspirations.

Literature Review

National Curriculum Framework for Physical and Health Education

The PHE curriculum under the KSSR and KSSM framework is structured around six fundamental pillars: communication, spiritual, attitude and values, humanity, personal skills, physical development and aesthetics, and science and technology. These pillars support the comprehensive development of students, emphasizing both physical and aesthetic growth crucial for personal well-being. The PHE secondary school curriculum adopts a diverse array of pedagogical strategies to facilitate meaningful learning experiences. These include mastery learning, inquiry-based learning, project-based learning, the teaching games for understanding approach, an integrated approach, contextual learning, and the application of multiple intelligences theory (Curriculum Development Division, 2017; 2018).

For the evaluation of these educational strategies, teachers are recommended to employ a combination of assessment methods. These methods range from informal assessments to formal evaluations using varied tools such as rubrics, portfolios, self-assessments, and traditional pen-and-paper tests. This blend of assessment techniques is designed to accurately gauge the students' levels of mastery across different dimensions of the curriculum, thereby ensuring a holistic educational approach that aligns with the framework's objectives (Curriculum Development Division, 2017).

The Physical Education curriculum aims to enable students to achieve the following objectives:

1. Explore concepts and principles in Physical Education through an inquiry approach.
2. Apply knowledge while participating in physical activities.
3. Use Physical Education terminology to communicate effectively.
4. Perform skills effectively in various contexts.
5. Demonstrate and apply various skills and techniques effectively while participating in physical activities.
6. Demonstrate and apply various principles, strategies, and concepts of movement and fitness during physical activities.
7. Collaborate and communicate effectively during physical activities.
8. Build positive relationships and exhibit social responsibility while participating in physical activities.
9. Display Physical Education learning experiences during physical activities.
10. Practice and maintain a healthy lifestyle for well-being (Malaysian Inspectorate Group, 2019).

The curriculum document known as the Curriculum and Assessment Standard Document delineates PHE content and its focus through Content Standards, Learning Standards and Performance Standards. The Content Standards consists of the main skills which students need to master, whereas the Learning Standards contains the breakdown of the Content Standards which serves as specific focus for instruction and assessment activities. The Performance Standard is established using a rubric that outlines the levels of mastery in skills and fitness achieved through the teaching and learning process using incremental progressions. The rubric is structured around cognitive, psychomotor, and affective taxonomies to evaluate achievement across these three domains. The Performance Standard includes six hierarchical Mastery Levels, with Level 1 indicating the lowest performance and Level 6 representing the highest. Each statement is presented generically to give a comprehensive overview of student progress and individual reporting (Curriculum Development Division, 2017; 2018).

Malaysian Physical and Health Education Teachers' Assessment Practice

Studies measuring Malaysian PHE teachers' issues and challenges in CBA implementation often involved examining the aspect from instructional perspectives (Zakaria et al., 2024). Sports equipment and facilities (Mohamed et al., 2019; Veloo & Ali, 2016; Veloo & Krishnasamy, 2017; Wee, 2019; Wee et al., 2020); teachers' content knowledge (Veloo & Ali (2016); Veloo & Krishnasamy (2017) and pedagogical content knowledge (PCK) (Veloo & Ali (2016); Wee et al (2020); student and parent factors (Kilue & Mohamad (2017) were salient themes in literature on issues and challenges to effective PE implementation. Administrators'

attitudes and perceived lack of support were also strongly acknowledged. School administrators were found to weight and prioritize core subjects, particularly subjects included in the Sijil Pelajaran Malaysia (SPM), over PHE Fonyi & Soon (2021); Wee (2017); less responsive in resolving issues pertaining to PHE teaching and learning improvements Veloo & Krishnasamy (2017); Wee et al (2020); and were more likely to assigned untrained teachers to teach PHE to complete the teachers' weekly loading hours (Wee & Chin, 2020).

Malaysian teachers' PE assessments have been reported to be largely at a low to moderate level of implementation (Abdullah et al., 2015; Yusoff et al., 2024; Saad et al., 2016; Veloo & Krishnasamy, 2017; Zakaria et al., 2024). Yusoff et al (2024) investigated the views of Malaysian PHE teachers on assessments, concentrating on four key areas: improving teaching and learning, school accountability, student accountability, and relevance. The study found that teachers regarded assessment as an important means to enhance the educational process. Nonetheless, they had reservations about its precision in evaluating student skills and performance. Lebar et al (2013); Veloo and Krishnasamy (2017) revealed PE teachers' low level of understanding of PE assessment and PHE curriculum assessment requirements, while Abdullah et al (2015) admitted that only a quarter of the PE teachers in their study were competent in test development. Zakaria et al (2024) shared that PE teachers demonstrated a moderate level of engagement with tasks related to traditional assessments; however, very few teachers in their study embarked on the test planning process, a finding similar to (Saad et al., 2016). Zakaria et al (2024) added that teachers' instructional practices were not evidence-based, with assessment information that could not be used to inform interventions and instructional improvements due to misalignment between the learning objectives and the implemented assessment tasks. Saad et al (2016) reported similar findings, noting the discrepancy between the curriculum expectations and the administered assessments. Teachers' adoption of psychomotor-focused assessment tasks was consistent with the intended learning objectives, however with limited opportunities for students to engage in assessments addressing the learning within affective domain (Zakaria et al., 2024).

Method

This study adopted an explanatory qualitative research design and employed in-depth interviews to explore causal relationships. In contrast to exploratory research, which primarily aims to uncover new insights into relatively unknown topics Cresswell & Poth (2016), explanatory qualitative research seeks to understand the underlying reasons for specific issues and challenges and to evaluate their impacts (Naess, 2018). This methodology focuses on identifying patterns and themes that elucidate the phenomena under investigation (Nascimento & Freitas, 2023). Whereas exploratory research typically establishes a foundation for future studies by generating hypotheses, explanatory qualitative research produces actionable insights that can inform interventions or improvements (Naess, 2018; Patton, 2014).

Within the scope of this study, the explanatory qualitative design facilitates a thorough exploration of the intricate and multifaceted nature of assessment practices in educational environments. This approach allows the researcher to probe deeply into the complexities of teachers' experiences and the contextual factors influencing their assessment strategies (Naess, 2018). By utilizing in-depth interviews, the research reveals the foundational reasons for specific practices, the challenges educators encounter, and the potential effects on student learning outcomes. This method provides a comprehensive understanding of the

dynamics involved, yielding critical insights that can guide targeted interventions and enhancements in assessment practices (Naess, 2018; Nascimento & Freitas, 2023).

The interview participants comprised four teachers, with two from Sabah and two from Sarawak, the two states located in the eastern part of Malaysia. The study employed purposive sampling through the use of a maximum variation sampling technique for participant selection. Purposive sampling involves selections based on the researcher's judgment to ensure the sample embodies characteristics pertinent to the study's objectives (Palinkas et al., 2015). This technique is widely recognized as effective in qualitative research for its ability to yield significant insights (Benoot et al., 2016; Palinkas et al., 2015). The adoption of maximum variation sampling, or heterogeneous sampling, facilitates the inclusion of participants who bring a broad array of perspectives due to their diverse experiences and backgrounds (Stewart, 2024). In this study, it enabled the gathering of a rich and varied collection of insights into the challenges, barriers, and opportunities associated with implementing assessment practices in physical and health education. Proponents of the purposive sampling method assert that the maximum variation technique is particularly suitable for studies aimed at identifying core and shared themes across a diverse sample, while also capturing unique, divergent insights (Alele & Malau-Aduli, 2023).

Several strategies were employed to ensure the validity and reliability of the interview process and the data obtained. The interview questions were first subjected to a pilot phase, involving a Physical and Health Education (PHE) teacher from an independent school, to assess the clarity, flow, and any potential ambiguities. Additionally, the accuracy and consistency of the interview transcriptions were meticulously verified during the transcription process, applying uniform coding criteria to ensure consistent data interpretation. Moreover, member checking was implemented, wherein the accuracy of the transcriptions was validated by sharing the transcribed texts with the interviewees for their feedback. Collectively, these measures significantly enhanced the robustness of the research methodology.

Data was analysed thematically using an inductive approach. Inductive analysis allows patterns, themes, and categories to emerge directly from the data, unimpeded by pre-existing assumptions and expectations. This method is widely regarded as suitable for thematic analysis in the social sciences, providing the flexibility to adapt as new themes arise (Thomas, 2006). The analysis followed the six stages proposed by Braun and Clarke (2006): familiarization with the data, generation of initial codes, searching for themes, reviewing themes, defining and naming themes, and generating the report endeavour.

Findings

The study aimed to answer two research questions: What were the issues and challenges that obstructed the effective implementation of valid and reliable PE assessments? And how did the persistent and unaddressed challenges influence the shaping of teachers' assessment norms and practices? The thematic analysis identified five primary themes of issues and challenges to effective PE assessments: the complexity of CBA, complexity of PE assessments, lack of standardized assessment approach, resource constraint, and administration-related issue. Additionally, two primary themes were extracted from the analysis involving assessment norms and practice: compromised integrity in scoring, and inconsistencies in assessment practice. These themes, along with their corresponding sub-themes, are explored in detail in the subsequent discussion.

Some of these key issues and challenges intersected and overlapped with one another due to the inherent complexity and interconnected nature of assessment, curriculum, and

pedagogical strategies. Nonetheless, for clarity the following discussions are structured around these themes as distinct categories. This approach allows for a detailed exploration of each theme, facilitating a more effective discussion of the underlying issues and norms.

Demographic Information of the Participants

The participants of the study include four PHE teachers, with two from Sabah and two from Sarawak, each representing different geographical areas within these states. To maintain confidentiality and for ease of discussion, pseudonyms are used for identification purposes. The demographic information of the research participants is presented in Table 1.

Table 1

Demographic information of research participants

Pseudonym	State	Age	Gender	Years of Teaching	Academic Background	School Type	School Location	Justifications for Selection
Teacher A	Sabah	30	Female	5 years	Sports Science	Trust School	(Remote) Sparsely populated area	Immediately appointed as PHE Head of Department upon first posting
Teacher B	Sabah	38	Male	15 years	Sports Science	Project School (Sports)	Urban area	The head of state coaches for a major sport and is currently teaching in a school championing the sports
Teacher C	Sarawak	41	Female	18 years	Sports Science	Daily secondary school	Semi-rural area	Class teacher, has taught PHE since first posting
Teacher D	Sarawak	42	Male	15 years	Sports Science	Daily secondary school	Urban area	An experienced teacher but only assigned with PHE

								loading 6 years ago
--	--	--	--	--	--	--	--	------------------------

Issues and Challenges

The first research question sought to identify issues and challenges in effective PE assessments. The analysis indicated that PE teachers viewed complex requirements and documentation demands of CBA as a major challenge. Teachers felt the need to document extensive evidence of student mastery across six competence levels overwhelming, especially given their heavy teaching and administrative loads and large class sizes. They believed the detailed nature of CBA was time-consuming and often impractical in real-time physical activity settings. Additionally, the perceived lack of standardized rubrics and structured scoring tools led to variability in assessments, further heightened by the isolation in which teachers operated, without coordinated or moderated assessment practices. Large class sizes, resource and staffing constraints, including the assignment of non-specialist PE teachers, further hindered effective assessments. The discussion below details these themes and sub-themes on issues and challenges faced by the PE teachers.

Requirements of Classroom-Based Assessment (CBA)

One of the issues that impeded teachers' PE assessments was the complex requirements of CBA assessments. Teachers found the assessment process within CBA complex and the documentation required was overwhelming them.

Complexity of Classroom-Based Assessment (CBA) Process

The teachers viewed the CBA assessment as complex due to the alignment between the content standards and learning standards, the recommended instructional strategies, formal and informal assessment activities that yield evidence for teachers to determine level of student mastery. In CBA, students' levels of mastery are mapped on six levels of increasing competence for any particular skill acquisition. Hence, teachers' assessments must consist of items or tasks that would yield evidence of learning leading to the placement of students on one of these levels. The teachers in the study, however, found such structure and assessment requirements to be complex and time laborious, particularly given the teaching and administrative load as well as large class size that they had to handle.

Teacher A admitted to feeling overwhelmed by the specific and detailed nature of CBA and the large class size she had. She added, not all students can meet the standards prescribed. Teachers A and B also associated the complexity of the assessment process with class size and teaching loads. This suggests that fewer teaching hours and smaller class sizes would enable teachers to dedicate more time and attention to carrying out the complex process of CBA. The detailed nature of CBA and the complexity of its assessments require significant understanding and skills, which can be challenging and time-consuming.

Extensive Documentation Requirements

All teachers in the study viewed the documentation process in CBA as extensive and, at times, could be overwhelming. Teacher C, for example, admitted to feeling burdened by the evidence and documentation requirements, particularly when assigning students to band 1 and band 6. The assessment process in CBA required tracking of student achievements for each PE skill taught, contributing to the assessment complexity. According to the teachers, this would require significant documentation and data entry, increasing the administrative

load on teachers. Additionally, student placement on the band progression would need to be supported by sufficient evidence, hence incurring the need to the teachers to provide sufficient documentation that supported their decisions on student placement.

Inherent Complexities of PE Assessments

All the research participants highlighted the focus on individual assessment in CBA as one of the main reasons for its complexity. Moreover, the focus on physical nature of the assessment activities meant that assessments often needed to be conducted in real-time, observing students as they performed various physical tasks. The teachers highlighted the difficulty in assessing individual performance when students were engaged in group activities. They noted that the complexity increased when all students participated simultaneously in group activities.

Teacher A highlighted the subjective challenges of CBA, particularly in PE, where individual performance sometimes required evaluation in a group setting. She believed that this situation complicated the assessment process because it could be difficult to isolate individual contributions and skills during group activities. The subjectivity arose from the need to interpret individual performance amidst collective participation, which can blur the lines of individual achievement and skill level.

Both Teachers B and C emphasized the practical difficulties in assessing individual student due to large class size and limited time, Teacher B, however, noted the advantage of real-time assessment in physical settings, where he could directly observe whether students were mastering the physical activities, aligning with the CBA's focus. However, he questioned the feasibility of such assessments given the time constraint and the overwhelmingly high number of students.

Lack of Standardized and Systematic Assessment Approach

The lack of a structured approach to measuring learning led to significant variability in PE assessments. The following discussion explores these themes and subthemes in detail.

Lack of Structured Approach to Measuring Learning at Individual and Group Levels

The thematic analysis highlighted variability in how opportunities were provided for students to demonstrate acquired skills, as well as in the execution of the scoring process by teachers. Three teachers in the study, Teachers B, C and D, appeared to follow a lesson format that was dynamic, with opportunity for assessment at individual and group levels. Of the three teachers, Teacher B and Teacher D were found to be more adaptive, flexible and provided students with sufficient opportunity to demonstrate the learned skills individually as well as in small groups and whole group structures. Teacher B, for example, placed emphasis on students' skill acquisitions in tandem with how the learned physical skills should be applied in real context. Hence, he incorporated group-based game in his lessons in evaluating student mastery to apply the skills in the authentic context.

Teacher A's description of her lessons, however, revealed minimal opportunities for students to showcase their skills at individual level. She began her lessons with explanation and demonstration of the intended skills before getting students executed the acquired skills at group level, often involving whole-class structure. According to Teacher A, "It's difficult to explain strategies for physical skills to students. We can tell them the strategy, but during the game, it seems like they forget they are learning because they are playing and enjoying themselves. So, it's hard for teachers to stop them to correct their strategy unless it's at the

end of the class. You've just played. If there's a moment, for example, if I find a suitable time when I see them applying a strategy, then I give them feedback, reminding them and saying, "You, you were doing this just now. Good job, you used the strategy that I taught." But if they are in the middle of playing, it's really impossible to stop them because they aren't focused. We summarize skill mastery during the cooldown at the end of the class." Her admission revealed lessons that were not well structured with instructional activities that were not facilitated accordingly. Provision of feedback, to Teacher A, could only be coordinated when she was presented with the opportunities to do so rather than a routine established from well-planned and structured lessons.

Lack of Standardized Scoring Tools

The teachers had differing opinions with regard to the adoption of CBA rubrics in scoring PE performance. Only Teacher D viewed the band levels as a form of rubric, the remaining three acknowledged the absence of standardized rubrics in CBA for assessing PE skills. When asked to describe CBA, Teacher D had likened CBA as an assessment instrument, On the other hand, Teacher C was quick to mention the lack of rubrics made available to teachers under CBA. Additionally, she was not convinced that the availability of scoring rubrics would improve the quality of her scoring.

The analysis on the adoption of CBA rubrics for scoring PE performance disclosed a divide among teachers. While Teacher D recognized the structured approach of CBA as beneficial for assessing student progress and mastery, the majority expressed concerns about the lack of standardized rubrics and the potential inconsistencies this caused in evaluations.

Siloed Assessment Practice

The analysis revealed specific instances that suggested the practice of siloed assessment, where individual teachers operated independently without a coordinated or standardized approach. Each teacher appeared to have their own approach to assessment, and there was no mention of a formal moderation process to align these assessments for consistency or to enhance inter-rater reliability. The analysis also highlighted an assessment practice that was guided primarily by personal judgment and the specific constraints of the classroom environment, resulting in varying assessment standards within the same school.

Responding to the interview question about whether the PHE teachers in the school had a system in place to ensure high scoring consistency across assessors and alignment of marking standards, all four teachers emphasized the isolated assessment practice. Two teachers acknowledged the variability in the scoring process that existed across teachers in their schools but did not seem to be aware of the importance of ensuring scoring consistency. All teachers held strongly to the belief that student assessments were the sole responsibility of the subject teachers, hence matters pertaining to it should not be discussed with other teachers.

Resource Constraints

The participants had expressed the constraints in their PE assessments due to the availability of insufficient sports equipment and less than ideal conditions of the sports facilities in the schools. Limited space was one of the main issues for Teacher B to carry out his PE assessments effectively. His secondary school is located in the same area with a pre-school and a primary school, with all the three schools sharing the same sports facilities. He

emphasized the necessity of larger spaces to effectively conduct assessments, highlighting the logistical constraints faced by the schools.

The reference towards the lack of sufficient sports equipment were made several times by all the teachers in different parts of the interview. Teacher C questioned the viability of completing all the skills in the syllabus when her school was not equipped with appropriate facilities that would enable certain skills to be taught. Teacher C said, "In PE we have skills that cannot be implemented due to problems with equipment, facilities and so on. So, that creates a problem here. How are we supposed to assess cricket, for example? There's cricket in the syllabus but we don't have cricket field. The same goes for tennis, we don't have tennis court here." Responding to the next question on how she managed to complete the syllabus given the lack of resources, Teacher C clarified, "I just focus on the theory [for the affected skills]."

Administration-related Issues

The implementation of CBA had been significantly influenced by administrative factors, particularly large class sizes, staffing constraints and the provision of trainings.

Large Class Size

The analysis highlighted large class size as one of the underlying factors contributing to most issues and challenges in effective CBA assessments. All the teachers had between 30 to 35 students in each class they taught. Teachers A, B, and C believed that this number was viewed as less than ideal for PE assessments, particularly when carrying out real-time assessments at the individual level involving psychomotor skills.

When asked about their class sizes, Teacher A indicated that she had about 30 to 40 students in a class; Teacher B admitting to teach nearly 10 classes with almost 400 students in total over 30 periods in a week; whilst Teacher C clarified that with seven classes and nearly 200 students in total, viewing assessment at individual student level as almost impossible. Responding to a question on what they believed to be a manageable number of students in a class for effective assessments, all three teachers agreed that 10 to 15 students would be an acceptable group size.

Staffing Constraints

The assignment of non-option PHE teachers due to staffing constraints was identified as one of the main reasons for high variability in PE assessments in schools involving two teachers in the study. Teacher B noted that teachers who were not specialized in PE often ended up teaching the subject due to scheduling or staffing needs: "What I mean is... these teachers are not PE teachers; they are assigned PE classes because their schedule still has 2 or 4 periods open." This variability was intensified by the lack of a standardized approach or consistent training across teachers, which is crucial for ensuring equitable and fair assessments.

In the case of Teacher A's school, a constantly changing timetable did not foster a sense of commitment among her non-option PHE colleagues to diligently carry out PE assessments. Teacher A explained that, during that particular year, teachers had been assigned and reassigned to different subjects at least four times. She clarified that the teachers' timetable would be reviewed and changed every time the school received new teachers. However, the constant changes in the timetable did not foster a sense of loyalty towards the subject. Citing an example, she explained how she and another colleague, who was an option PHE teacher, had carried out discussions and even demonstrated certain skills. Their enthusiasm was not

well-matched, as other teachers were not interested because they knew they would not be teaching the subject for long.

Lack of Training and In-School Training System

Part of the variability in assessment practice was attributed to the lack of clarity among teachers regarding effective PE assessment practices and the varied pedagogical skills necessary to support valid and reliable assessments. This was predominantly due to the absence of structured training systems in the participants' respective schools, leading to inconsistencies in assessment practices across teachers. Among the four participants in the study, only Teacher D was able to answer knowledge questions related to CBA, while the remaining three teachers struggled to explain the difference between content standards and learning standards, distinguish between DSKP and CBA, and comprehensively describe the levels of mastery. Additionally, the teachers exhibited confusion over the need to assess physical skills and believed that summative assessments would be sufficient to measure these skills. When asked to rate their own understanding of CBA, Teacher B candidly admitted that he grasped only about half of what the component entails.

The findings showed that training initiatives were only made available to new teachers and its availability was dependent on the PE's head of department discretion. In Teacher A, B and C's schools, new teachers were expected to consult senior teachers in obtaining information related to teaching and learning with no structured system, such as mentoring, peer teaching and teaching observation, in place in all the participants' schools. Only one school, the school of Teacher D, had a systematic training structure in place. New teachers were provided briefing and were shown the correct ways in teaching and assessing PE through peer learning. However, all the teachers in the study admitted to not receiving any form of CBA-related course or workshop at least in the past eight years.

Assessment Norms and Practice

The prolonged and unaddressed exposure to various issues and challenges, coupled with the lack of appropriate interventions and support for PE teachers, have inadvertently solidified certain assessment practices into entrenched norms. The second research question sought to identify the effects of these unaddressed challenges on teachers' assessment norms and practices. Prolonged exposure to the challenges identified as the findings for the first research question, without adequate support or intervention, had forced teachers to adapt by simplifying assessment criteria, often at the expense of assessment integrity. This adaptation includes the practice of grade inflation and averaging scores. Such practices not only questioned the validity and reliability of assessments but also fostered an environment where subjective judgments prevailed over standardized criteria. Additionally, inconsistencies in assessment practices arose from the isolated responsibility of individual teachers and a lack of standardized scoring tools. Resource limitations forced teachers to modify the syllabus, delivery methods, and assessment techniques, often shifting focus from psychomotor to theoretical assessments. These adaptations, while practical, resulted in variability and misalignment with the overarching educational goals.

Grade Inflation

The mismatch between administration's expectations and actual student performance created a dilemma for teachers who felt compelled to inflate grades to meet these expectations, resulting in potential inaccuracies of student assessments. The responses

addressing the practice of grade inflation were noted from three teachers in the study, largely involving students who demonstrated learning at lower band levels.

The three teachers communicated the school expectation for all students to be placed at band 3 following lessons and instructional activities. Teacher B found that it was easier to accurately assess excellent students but there remained a challenge in objectively measuring less obvious performance levels, particularly those at lower band levels. For Teacher C, the issue involved students who were at the extremes on both ends. Teacher A explained that not all students participated in physical activities. Despite assessment indicating some students should be at band 1, schools would question these placements, pushing teachers to raise grades to the second or third level. This pressure made grading seemed arbitrary and driven by external expectations rather than student performance.

Assessment Average and Practical Adjustments

Teacher A, B and C explicitly addressed the practice of averaging scores to overcome practical assessment challenges due to large class size and in coping with administrative expectations. To Teacher A, the practice of averaging scores came about from the limitations in evaluating individual skills during sports activities, particularly due to the practical constraints of the teaching environment. When dealing with large groups of students, Teacher B mentioned the necessity of averaging out performances due to the sheer number of students, which did not allow for a detailed assessment of each individual. Teacher C's approach involved categorizing students into three ability levels (low, moderate and high ability) in the execution of the taught skills and then averaging the scores within these categories to assign a representative score. Although this method acknowledged individual differences to some extent, it still relied on averaging to manage the practicalities of assessing individual students.

While averaging scores facilitated the management of large classes and aligned with administrative requirements, it potentially oversimplified the diverse capabilities of students, leading to less personalized educational feedback and potentially masking the true performance levels of individual students.

Modifications Made to Syllabus Structure, Delivery and Assessments

The analysis revealed that prolonged exposure to resource limitations compelled teachers to restructure the syllabus, alter the delivery methods, assessment techniques, and focus. When asked how resource limitations affected her PE assessments, Teacher A replied, *"... we can teach football, but there aren't enough balls. We end up having to use whatever balls are available for the students to kick around. We do manage to conduct the lessons, but teachers need to be creative. Sometimes they even have to kick a volleyball. That's the problem with the lack of proper equipment."* Teacher A's response highlighted the need for teachers to be highly creative and adaptive to enable the intended skills to be taught and mastered despite the absence of necessary equipment.

For Teacher A, resource limitations were a major issue. One of the strategies adopted for coping with the lack of proper equipment was alternating the topics and skills between teachers at her school, enabling them to rotate equipment. However, this resulted in low uniformity in teaching and assessment, as well as disjointedness in the flow of intended skills outlined in the curriculum. Sometimes, students were taught more advanced skills when equipment and tools for much simpler skills were being used in other classes. This disrupted the natural progression of skill acquisition, as students were forced to jump back and forth between learning one skill and another. At Teacher B's school, one field had to be shared with

two other schools located within the same compound. The small size of the field, combined with the need to conduct PE assessments while the field was also in use by students from other schools, limited the assessment effectiveness of the PE teachers.

In addition to variability in assessment practices, prolonged exposure resource constraints compelled the teachers to revise the skills emphasized during PE. For instance, at Teacher C's school, teachers shifted their focus to theoretical knowledge and understanding for skills that lacked the necessary equipment, necessitating the adoption of pen-and-paper-based assessments. This practice led to a significant misalignment with the overarching educational objectives, as it redirected the focus from physical activity to a cognitive understanding of the activity.

Discussion

The study provided valuable insights into the assessment challenges faced by PE teachers as they highlighted the deep-seated issues that require urgent attention and intervention. While the limited number of participants does not allow for the findings to be generalized across all educational contexts, the insights gleaned are vital for understanding the real-world implications of policy decisions and educational practices on PE teachers' day-to-day professional lives.

Fundamental Gaps in Knowledge and Understanding

The findings illuminated gaps in teachers' knowledge about CBA, teachers' PCK, and teachers' assessment competencies. The necessity to conduct individual assessments in real-time required a sophisticated set of skills that teachers in the study found themselves under-equipped to handle. PE teachers also encountered difficulties with the detailed and rigorous requirements of CBA, demonstrating partial understanding of what CBA entailed, its aspirations, and the necessity for its implementation. For instance, teachers expressed difficulty with the detailed demands of CBA, finding these overwhelming due to their complexity and the time required for proper implementation. These findings resonate with the findings of Abdul Razak et al (2023); Tan et al (2022); Veloo (2016) who shared teachers viewed CBA as administratively burdensome and detracted them from teaching time, leading to a decrease in overall instructional quality.

Interwoven with the need for deeper assessment knowledge was the challenge of applying these skills effectively within the pedagogical setting. Teachers reported that the dynamic nature of PE, requiring real-time, personalized feedback during physical activities, presented practical challenges. Such findings are not uncommon with studies by Veloo (2016); Veloo and Krishnasamy (2016); Yusoff et al (2024); Wee (2017, 2019); Wee and Chin (2020); Wee et al (2021) reporting that teachers struggled to facilitate effectively the dynamics of PE lessons. CBA implementation was also reported to be at low level (Abdullah et al., 2015; Lebar et al., 2013; Veloo & Krishnasamy, 2017). Lessons were characterized by ineffective classroom management strategies (Veloo, 2016; Veloo & Ali, 2016); with PE teachers employing lowly innovative instructional approaches (Zakaria et al., 2024); unable to identify learning weaknesses (Mat Yusoff et al., 2024; Wee, 2014), lacking the provision of constructive and targeted feedback (Wee, 2014; 2020), with lessons consisted of missed opportunities for meaningful feedback in correcting errors and learning weaknesses (Zakaria et al., 2024). Additionally, PE teachers' assessment competencies were reported to be at low level (Abdullah et al., 2015; Lebar et al., 2013; Saad et al., 2016), with misalignment in assessment task with curricula and policies (Dyson et al., 2018; Zakaria et al., 2024). Zakaria et al (2024)

disclosed that PE assessment practice was a routine task with no efforts being made to link evidence of learning to improve instruction or implement intervention strategies.

The teachers in the study viewed 30 students in a classroom as particularly too many for them to carry out effective PE lessons, particularly considering the need for individualized real-time assessments. However, 30 students in a group are commonly viewed as an acceptable size within the landscape of Malaysian education, where most schools in urban areas are overcrowded with typically more than 40 students per classroom (Abdul Majid et al., 2014; Rasyidah & Wafa, 2020). Hence, this revelation pointed to the significance of PCK and assessment skills in supporting effective instructional and assessment activities.

The Absence of the Culture of Sharing and Collaboration

A critical concern that emerges from the study is the individualized approach to assessment practice. This approach led to substantial variability in how assessments were conducted, with each teacher relying predominantly on their individual judgment and preferred methods. This fragmented approach not only resulted in inconsistencies in how student performances were evaluated across different classes but also diminished the potential for sharing best practices and learning from collective experiences (Zakaria & Ab. Wahid, 2023; Zakaria et al., 2016). The findings indicated that there was little to no discussion among teachers regarding scoring standards, which widened the gap in scoring consistency. The attitude that each student's performance was solely the subject teacher's responsibility further isolated teachers, preventing a unified approach to assessment that could benefit from shared insights and strategies (DeLuca et al., 2019).

Silo mentality, a term commonly associated with team-based efforts, collaboration, and professional development, has significant implications for both individual practitioners and organizations (Neill & Jiang, 2017). This isolationist approach hinders the creation of a cohesive assessment strategy that could leverage the collective expertise and insights of multiple educators. Waal et al. (2018) emphasize that such a mentality negatively impacts organizational performance, innovation, and openness to change.

Central to shifting teachers' siloed assessment practices is increasing their knowledge and awareness of the importance of assessment consistency, particularly in ensuring fairness and equity among PE teachers involved in the same assessment. Teachers need to understand that discussions between colleagues are not a means of transferring responsibilities to others. By working together, teachers can develop a more coherent and consistent approach to evaluating student performance.

Development of Compromised Scoring Practice

Over time, the relentless pressures of complex assessments, combined with inadequate support and training, have led to the emergence of compromised norms in assessment practices that may not align with educational best practices. For instance, grade inflation and score averaging, as noted in the findings, have become prevalent norms among the teachers in the study. Administrative pressure is cited as a common cause for educators to engage in such practices (Chowdury, 2018; Boleslavsky & Cotton, 2015; Griffith & Sovero, 2021).

One of the most pressing concerns with grade inflation is its impact on educational equity. When grades are inflated, it becomes difficult to identify students who may be struggling or those who genuinely excel, thereby denying them necessary support or recognition (Chowdury, 2018; Griffith & Sovero, 2021). Averaging scores can mask the diversity of individual student performances, potentially leading to evaluations that do not

accurately reflect each student's abilities or progress (Anderson, 2018; Doz, 2023). Both methods of scoring can perpetuate a cycle of mediocrity, where high-performing students are not sufficiently challenged, and lower-performing students do not receive the intervention they need (Boring & Philippe, 2019; El-Alayli et al., 2018). Over time, this can widen the achievement gap, contrary to the goals of equitable education systems (Anderson, 2018; Ehlers & Schwager, 2016; Griffith & Sovero, 2021).

Low Prioritization of Physical Education

The study reveals fragmented assessment practices, resource constraints, and administrative inefficiencies as areas that could potentially be improved if the school administration placed greater priority on the importance of quality PE instruction and assessment. The lack of a structured assessment system, collaborative culture, and adequate training fostered teachers' adoption of isolated assessment practices. The absence of a formal moderation process between PE teachers to ensure scoring consistency, and the lack of support mechanisms such as peer teaching and mentoring, reflect the need for effective organizational norms to be established at the school level.

Lack of sports equipment and facilities is a long-standing issue. Although it is reasonable to expect PE teachers to be adaptive and innovative in the absence of certain types of sports tools and equipment, such measures should only be temporary. While expecting schools to be provided with facilities such as tennis and cricket courts would require more than budget planning and expenditure at the school level, the gradual purchasing of sports equipment is something that schools can actively pursue. Students need opportunities to learn and demonstrate their skills within authentic settings, not through improvised means.

Administrator attitude is an underlying theme in various studies that investigated issues associated with PE instructional and assessment process (Wee, 2017; 2019; Veloo, 2016; Veloo & Ali, 2016). Similar to the findings of this study, Wee (2019) shared that administrators in his study did not discuss PE teaching assignments with teachers, with a majority of the teachers admitting that the assignment was done without the consideration on their specialized area. Wee and Chin (2020); Kamaruzaman et al (2014) made similar assertion having found that school administrators were more likely to assign untrained teachers to teach PHE to complete the teachers' weekly loading hours. Additionally, school administrators were reported to place greater value on subjects included in high-stakes examinations over those that are not (Fonyi & Soon, 2021; Wee, 2017; 2019). Consequently, PHE lesson hours have been sacrificed for the learning of other subjects (Veloo & Md Ali, 2016; Wee et al., 2020).

Implications, Recommendations and Future Study

The findings from the study illustrate the urgent need for robust professional development (PD) programs specifically tailored to enhance PE teachers' proficiency in CBA, sound assessment knowledge and skills, and varied pedagogical approaches and strategies that support the unique nature of PE assessments. PD programs should equip teachers with strategies to seamlessly integrate various forms of assessment. The study reflected on the real-time challenges of assessing diverse student abilities within large classes, aggravated by limited resources and inadequate sports equipment and facilities. Hence, PD initiatives should include components that specifically address resource limitation challenges. Training teachers on how to effectively utilize limited resources while still adhering to assessment standards can mitigate some of the adverse effects identified.

The findings revealed significant variability in assessment practices, largely attributed to the isolated manner in which PE assessments were conducted. This isolation led to inconsistencies across evaluations. The formation of learning communities within schools, such as Professional Learning Teams (PLTs), mentoring programs and peer teaching initiative, are some of the practical strategies that school leaders and administrators can undertake. The district education department and schools should actively work towards bolstering the lifelong learning ethos. This institutional support is crucial in signalling to teachers the value placed on continuous growth and learning. At the individual level, teachers themselves should be proactive in cultivating a lifelong learning culture and engaging in reflective practices. Reflective practices, such as maintaining journals, participating in peer observations, and engaging in feedback sessions, are essential tools that enable teachers to critically evaluate and enhance their instructional techniques and assessment methodologies.

Future studies should involve a larger and more diverse group of PE teachers across different regions. Conducting comparative studies between teachers in different educational settings, for example, urban versus rural schools, could illuminate how contextual factors influence assessment challenges and the formation of compromised norms. Engagement in longitudinal research could also offer deeper insights into the long-term effects of current assessment practices on both teachers and students. Additionally, studies with experimental designs aimed at measuring specific interventions to improve assessment practices would provide more robust evidence.

Limitations

While the study provided valuable insights into the assessment challenges and the emergence of compromised norms faced by PE teachers, it is important to acknowledge certain limitations that may affect the generalizability and depth of the findings. Firstly, the qualitative research design, while robust in exploring the experiences of educators, involved only four participants. This small sample size, confined to two specific regions of Malaysia, limits the ability to generalize the results across the broader educational landscape. The explanatory nature of qualitative research offers in-depth perspectives but also means that the findings are highly contextual and may not reflect the experiences of PE teachers in different or more varied educational settings. Additionally, the study's reliance on in-depth interviews, although effective in gathering rich, detailed data, might lead to a form of response bias, where participants provide socially desirable answers or reflect disproportionately on negative experiences. These limitations suggest a need for further research involving a larger, more diverse sample and potentially incorporating quantitative methods to provide a more comprehensive analysis of the prevalent assessment challenges.

Conclusion

In conclusion, this study has unveiled significant challenges faced by four teachers in the region of Sabah and Sarawak concerning PE assessments. The findings highlight a critical need for interventions that address the complexities of assessment practices and the professional development of teachers. The isolation in assessment processes and the lack of resources and support have contributed to the emergence of compromised assessment norms among PE teachers, impacting the quality of their evaluations. As such, this paper advocates for targeted professional development programs tailored to enhance teachers' assessment competencies and for systemic changes that foster a collaborative, resourceful educational environment. Additionally, the establishment of learning community and mentoring programs within

schools could significantly aid in standardizing assessment practices and enhancing the professional growth of teachers. Such initiatives are essential not only for improving the quality of PE assessments but also for ensuring that educational outcomes are equitable and conducive to fostering a lifelong learning culture among educators. Moving forward, it is crucial that these insights catalyse policy revisions and practical changes in school settings to better support teachers and students alike in the dynamic field of physical education.

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

- Abdul Majid, A. H., Hashim, S. H., Zain, M. R., Sarna, A. S., Lye, G. P., Chang, K. S., & Saran, M. (2014). Classroom management practices: Observations in selected Malaysian classrooms. *IOSR Journal of Humanities and Social Sciences*, 19(11), 54-58.
- Abdul Razak, R., Yusoff, M. S., Chin, H. L., & Marzaini, A. F. M. (2023). Evaluating teachers' pedagogical content knowledge in implementing classroom-based assessment: A case study among ESL secondary school teachers in Selangor, Malaysia. *PLOS One*.
- 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>
- Alele, F., & Malau-Aduli, B. (2023). *An introduction to research methods for undergraduate health profession students*. Pressbooks.
- Anderson, L. W. (2018). A critique of grading: Policies, practices and technical matters. *Education Policy Analysis Archives*, 26(49). <https://doi.org/10.14507/epaa.26.3814>
- Benoot, C., Hannes, K., & Bilsen, J. (2016). The use of purposeful sampling in a qualitative evidence synthesis: A worked example on sexual adjustment to a cancer trajectory. *Technical Advance*, 16(21), 1-12.
- Boleslavsky, R. & Cotton, C. (2015). Grading standards and education quality. *American Economic Journal: Microeconomics*, 7(2), 248-279.
- Boring, A., & Philippe, A. (2019). *In reducing discrimination in the field: Evidence from an awareness raising intervention targeting gender biases in student evaluation of teaching*. Working Paper.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp0630a>
- Chowdury, F. (2018). Grade inflation: Causes, consequences and cure. *Journal of Education and Learning*, 7(6), 86-92.
- Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications.
- Cui, Y. H., & Lei, H. (2015). Constructing a three-factor model of teaching-learning-assessment alignment. *Journal of East China Normal University*. 33(04), 15-22.
- 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.

- DeLuca, C., Chapman-Chin, A., & Klinger, D. A. (2019). Toward a teacher professional learning continuum in assessment for learning. *Educational Assessment*, 24(4), 267–285.
- Doz, D. (2023). Factors influencing teachers' grading standards in mathematics. *Oxford Review of Education*, 49(6), 819-837.
- Dyson, B., Cowan, J., Gordon, B., Po, W.D. & Shulruf, B. (2018). Physical education in Aotearoa New Zealand primary schools: Teachers' perceptions and policy implications. *European Physical Education Review*, 24(4), 467-486.
- Ehlers, T., & Schwager, R. (2016). Honest grading, grade inflation and reputation. *CESifo Economic Studies*, 62(3), 506-521. <https://doi.org/10.1093/cesifo/ifv022>
- Funyi, L., & Soon, C.C. (2021). Administrators' attitude towards the implementation of physical education in Selangor primary schools. *Pertanika Journal Social Sciences and Humanities*, 29(1), 291-309. <https://doi.org/10.47836/pjssh.29.s1.16>
- Gleddie, D. L., & Morgan, A. (2021). Physical literacy praxis: A theoretical framework for transformative physical education. *Prospects*, 50(1-2), 31-53.
- Griffith, A. L., & Sovero, V. (2021). Under pressure: How faculty gender and contract uncertainty impact students' grades. *Economics of Education Review*, 83.
- 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*. Tanjong Malim, Malaysia.
- Leong, N., & Chee, C. S. (2021). Administrators' attitude towards the implementation of physical education in Selangor primary schools. *Pertanika Journal of Social Sciences and Humanities*, 29(31), 291-309. <https://doi.org/10.47836/pjssh.29.S1.16>
- Malaysian Inspectorate Group. (2019). *Physical education health education teaching guide book*. Ministry of Education Malaysia.
- Yusoff, M. S., Chin, H. L., Abdul Razak, R., Basal, M. H., & Marzaini, A. F. M. (2024). Conceptions of assessment: Perceptions of physical and health education teachers in Selangor, Malaysia. *Pertanika Journal of Social Sciences and Humanities*, 32(1), 189-216.
- 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.
- Moh'd, S. S., Uwamahoro, J., Joachim, N., & Orodho, J. A. (2021). Assessing the level of secondary mathematics teachers' pedagogical content knowledge. *Eurasia Journal of Mathematics, Science and Technology Education*, 17(6).
- Naess, P. (2018). Validating explanatory qualitative research: Enhancing the interpretation of interviews in urban planning and transportation research. *Applied Mobilities*, 5(2), 186-205.
- Nascimento, E. A. D., & Freitas, J. S. (2023). Capturing causal complexity: A formal qualitative approach to explanatory research [Paper presentation]. *SEMEAD2023*. Indicado à premiação na área de Ensino e Pesquisa em Administração, Sao Paulo, Brazil.
- Neill, M. S., & Jiang, H. (2017). Functional silos, integration and encroachment in internal communication. *Public Relations Review*, 43(4), 850-862.
- Palinkas, J., Horwitz, S., Green, C., Wisdom, J.P., Duan, N., & Hoagwood, K. (2015). Purposive sampling for qualitative data collection and analysis in mixed method implementation research. *Administrative Policy Mental Health*, 42(5), 533-544.

- Patton, M. Q. (2014). *Qualitative research and evaluation methods: Integrating theory and practice*. Sage Publications.
- Rasyidah, G., & Wafa, S.W. (2020). A quantitative assessment of school environment as viewed by teachers. *Heliyon*, 6. <https://doi.org/10.1016/j.heliyon.2020.e04059>
- 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.
- Stewart, L. (2024). *Purposive sampling in qualitative research*. Atlas.ti Newsletter. <https://atlasti.com/research-hub/purposive-sampling>
- Kamaruzaman, S. A., Zaki, M., Che Hassan, & Julismah, Jani. (2014). Efikasi sendiri guru pendidikan jasmani terhadap pelaksanaan pengajaran mata pelajaran Pendidikan Jasmani. *Jurnal Kurikulum dan Pengajaran Asia Pasifik*. 1, 43-51.
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237-246.
- 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., & 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>
- 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.
- Waal, A. D., Weaver, M., Day, T., & Heijden, B. V. D. (2019), Silo-busting: overcoming the greatest threat to organizational performance. *Sustainability*, 11(23), 1-21.
- Wee, E. H. (2017). Quality control in physical education in Malaysia: Relooking at the national strategy for quality physical education. In G. Teh & S. Choy (Eds.), *Empowering 21st century learners through holistic and enterprising learning* (pp. 197–209). Springer.
- 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., & 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.
- 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.
- Zakaria, Z., & Ab. Wahid, N. T. (2023). Integrated assessment initiative at the forefront of curriculum, delivery and assessment efficiency. *International Putra InnoCreative Carnival in Teaching and Learning 2023 [Poster Abstract]*. Universiti Putra Malaysia.
- Zakaria, Z., Rizal, A. I. S., & Mohamed, A. M. D. (2024). Assessment in action: Investigating the practices of Malaysian physical and health education teachers. *International Journal of Academic Research in Progressive Education and Development*, 13(1), 2187-2204.
- Zakaria, Z., Care, E., & Griffin, P. (2016). Scaffolding instruction where it matters: Teachers' shift from deficit approach to developmental model of learning. *Journal of Education and Practice*, 7(23), 144-152.
- Zhao, L., Zhao, B., & Li, C. (2023). Alignment analysis of teaching–learning–assessment within the classroom: How teachers implement project-based learning under the curriculum standards. *Disciplinary and Interdisciplinary Science Education Research*, 5(13), 1-23.