# Examining South Korea's Elementary Physical Education Performance Assessment Using Assessment Literacy Perspectives

Yongnam Park [a]

## Abstract

This study examines the issues pertaining to South Korea's elementary physical education (PE) performance assessment, using an assessment literacy (Hay & Penney, 2013) perspective to propose future directions. Eight elementary teachers currently teaching PE were selected as participants. Data were collected through semi-structured in-depth interviews and on-site data analysis, and analyzed based on an inductive categorical analysis, the results of which were elicited from the four concept factors of assessment literacy. Four themes presented themselves during the research: first, teachers were unclear about the concept of assessment, making it difficult for them to carry out assessments effectively; second, assessments were conducted in a labor exchange and recycling manner, reducing their effectiveness; third, there was a lack of feedback; and fourth, there were teachers' critical thinking without pedagogy. These are the main problems in assessing PE in elementary schools. As for future directions, this study proposes the need for diversifying current teachers' education geared towards enhancing their assessment literacy capability, for providing on-site guidance to build students' assessment literacy, and for evaluating the assessment procedure.

**Keywords:** Assessment, Assessment literacy, Physical education assessment, Performance assessment.

## Introduction

As accountability in education is being increasingly emphasized, more importance has been attached to assessment, whether at the national, regional, or school level (Hardman & Marshall, 2000). South Korea has been highly ranked each year in the OECD-led Programme for International Student Assessment (PISA), and it has made assessments the focus of its education system to the extent that it is sometimes called "the Evaluation Republic." Subject evaluation activities are used to determine teaching quality; not only teachers, but also the students and their parents are evaluators.

In South Korea's evaluation-oriented school education, however, PE has been ignored and is in fact classified as optional; it is not considered for college entrance exams, and does not lend itself to performance evaluations, because of its unique characteristics. In this respect, standards for PE achievements and evaluations were promulgated, and research on developing an evaluation system was conducted at a national level, led by the Korea Institute for Curriculum and Evaluation (KICE). Performance assessment is an evaluation method that observes and judges students' knowledge, function and attitude in various authentic tasks and situations (McMillan, 2007). During the performance assessment process, teachers make comprehensive judgments through students' answers, outputs, and behaviors through several measurements and observations. Even

with such national efforts, teachers still face difficulties with the evaluation of PE classes in their teaching activities (Yoo, 2005).

These difficulties have also been reported outside Korea. Issues such as teachers' indifference in assessment, their lack of knowledge about assessment and their attitude in failing to recognize assessment as part of PE have consistently been raised in various PE assessment-specific studies (Annerstedt & Larsson, 2010; Hay & Penney, 2013; Matanin & Tennehill, 1994). In particular, elementary school teachers assigned to teach various subjects are experiencing relatively more difficulties in PE assessment, and it is vital to galvanize research and interest in this area. Studies on the extent of teachers' knowledge regarding assessment and how well they can be implemented and interpreted also need to be conducted (Thompson & Penney, 2015).

Various levels of recommendations thus far made regarding PE evaluations have been fragmented, and do not provide clear indications about the knowledge, techniques, abilities, and attitudes that are necessary for appropriately evaluating PE. By positioning PE assessments into social and cultural activities instead of treating them as measurement-oriented assessments from a scientific perspective, the assessment literacy concept recently proposed by Hay and Penney (2013) has garnered attention; it provides an alternate perspective, incorporating the knowledge, techniques, and attitudes

[a] Correspondence details: Yongnam Park, Department of Physical Education, Seoul National University, KOREA 71, 1 Gwanak-ro, Seoul, 08826, KOREA. Phone: (+82) 10-2976-5845. E-mail: mri4558@snu.ac.kr

contained in the assessment.

Assessment literacy generally refers to an ability to understand and utilize the information obtained from developing and evaluating assessment standards. This is indispensable for appraising teachers' understanding of their assessment procedures and assignments, and the quality of students' performance (Fullan, 2002). In addition, the PE assessment literacy concept proposed by Hay and Penney (2013) emphasizes understanding assessment procedures and being aware of both the efforts involved in requesting assessments and responding to the assignments, and the non-educational intentions and latent outcomes implied in the assessment procedures. Assessment literacy comprises four concepts: assessment comprehension, assessment application, assessment interpretation, and critical engagement with assessment. Assessment comprehension is related to teachers' knowledge and understanding of the conditions for achievement standards or efficient evaluations; assessment application concerns conducting assessments (including both teacher evaluation and student evaluation); assessment interpretation considers the social roles and interrelations of evaluations; and critical engagement with assessment refers to the capability needed for a natural assessment plan, implementation, and results, while recognizing the outcomes and influences of the assessments. The ideas behind assessment literacy represent a holistic perspective as to what teachers are aware of and how they conduct,

interpret, critique, and conceptualize assessments in sociocultural contexts; this helps foster a holistic assessment instead of an ad-hoc approach to individual issues. As seen in the analysis of the four conceptual factors of PE assessment literacy, they constitute a theoretical framework for elucidating teachers' expertise on assessments and the process of their evaluations; performance assessment can be conducive to exposing issues in PE assessments since they are a representative evaluation method. Therefore, this study examines the performance assessment issues in elementary school PE, with the intent of providing guidelines for future training for effective assessment.

## Material & Methodology

### Participants

The research participants were selected via the purposeful sampling method (Creswell, 2009). First, an initial pool was created with full-time teachers having five or more years of experience in elementary school teaching in South Gyeongsang Province, who were recommended by their colleagues. Second, the teachers from the initial pool who met certain specified criteria were selected as final research participants. These criteria were – the teachers should currently be teaching PE and conducting assessments as lead or full-time teachers, and have more than three years of experience in PE teaching and assessments. Eight teachers were selected, and their specific backgrounds are provided in Table 1.

**Table 1.** *Background information on participants*

| Pseudonym | Age / Gender | Teaching career | Teaching type | Experience in PE assessments | Grade level | Region |
|---|---|---|---|---|---|---|
| Shin-young | 30 / F | 7 years | Classroom teacher | 5 years | 6th | City |
| Min-jeong | 29 / F | 6 years | Classroom teacher | 5 years | 5th | Rural |
| Tae-ho | 29 / M | 4 years | Specialist | 4 years | 4-6th | Rural |
| Ha-jin | 41 / F | 17 years | Classroom teacher | 12 years | 5th | Rural |
| Jong-soo | 37 / M | 11 years | Classroom teacher | 10 years | 6th | City |
| Jae-in | 37 / M | 11 years | Classroom teacher | 8 years | 6th | City |
| Woo-jin | 33 / M | 8 years | Specialist | 8 years | 5-6th | City |
| Soo-won | 30 / M | 5 years | Specialist | 5 years | 5-6th | City |

### Data Collection and Analysis

On-site data were collected, including with in-depth interviews, documents, photos, and video files. In-depth interviews using semi-structured questionnaires initially lasted 60 minutes. The following questions were contained in the semi-structured questionnaire, which covered participants' specific tasks, their knowledge of PE performance assessment and their source of information, and their critical thinking about PE performance assessment: "1. What kinds of tasks do you specifically conduct during your PE performance assessment?"; "2. What efforts have you been making to enhance your

expertise regarding PE assessment?"; "3. What prerequisites would your ideal PE performance assessment need?"; and "4. What difficulties have you been through during your assessment of PE?" After the initial coding, additional 40-minute interviews were conducted with four research participants whose responses generated meaningful codes relevant to the four assessment literacy concepts. The additional interviews were based on detailed questions that examined the research participants' impartiality in the performance assessments, application of the performance assessment outcome, and teachers' efforts to develop students' assessment literacy.

As for on-site documents, data on the performance evaluation process for PE that research participants had conducted were elicited from the relevant documents and the resulting reports from previous PE performance assessments. Furthermore, the photos and video footage filmed during the participants' assessment of their students' PE performance were utilized since those data, though limited and indirectly confirming their on-site assessment, could be compared with what was stated in interviews.

The collected database was analyzed in the following stages, pursuant to an inductive categorical analysis as put forth by Patton (2002). By comparatively analyzing the collected on-site documents including research participants' plans for PE performance assessments, photos, and video footage, we obtained a comprehensive understanding of the PE performance assessment process. The theoretical framework consisting of the four conceptual factors of PE assessment literacy (Hay and Penney, 2013) was excellent in initiating a "start list" of pre-set codes (often referred to as "a priori codes"). However, the possibility of deriving other codes was not precluded in the coding process. Two coders worked independently for the initial coding, subsequently working on discordant codes again until they reached an agreement. The data analysis of in-depth interviews started with writing analytic memos after transcribing the interviews on the same day, from which initial codes were created by collecting repeated ideas during the segmenting task process. Thereafter, in-depth coding was conducted to create new inclusive codes by repetitively re-categorizing the initial codes. Finally, with a focus on the four conceptual factors of assessment literacy, outcomes were determined by inducing the themes that were interconnected among the created codes.

Using peer debriefing, we shared the initial data analysis with research participants to enhance the research veracity, performed checks to ensure data integrity, and also shared the entire research process with a university professor who had majored in Sports Pedagogy as well as three on-site teachers to guard against any error in the research process (Creswell, 2009). The entire process took place in accordance with the permission and regulations of the institutional review board (IRB) of the university with which the researcher of this study is affiliated.

### Results and Discussion

Through the data analysis process, four themes were constructed that aligned with the four conceptual factors of PE assessment literacy (Hay and Penney, 2013): (i) lack of clarity about the concept of assessment, (ii) labor exchange and recycling, (iii) a lack of feedback, and (iv) teachers' critical thinking without pedagogy.

*Assessment Comprehension: Lack of clarity about the concept of assessment*

Assessment comprehension signifies teachers' knowledge and understanding about assessments in educational contexts. As a fundamental premise for assessment efficacy (Hay & Penney, 2009), assessment comprehension refers to how much an individual teacher

is equipped with expertise and understanding of assessments and how they decide on the overall process and results of assessments; it is a starting point that determines the extent of influence the assessment results have on students.

Sufficient understanding of assessments should be preceded by an understanding of the concept of "assessment." However, the research participants exhibited confusion about assessment, evaluation, measurement, and their actual realization. Regarding this, Shin-young, one of the participants, stated:

> "Frankly, I am not quite sure what 'assessment' means. I am more familiar with the word 'evaluation,' and yet, at some point, the word 'assessment' started being used as well. As a teacher, I think PE performance assessments are conducted to evaluate, and classify into grades, students' physical functions and affective aspects related to classes. Is there much difference between 'evaluation' and 'assessment'?" (Shin-young)

What Shin-young stated regarding the conceptual definition of performance assessment still borders on a measurement-oriented assessment, which has long been conducted in PE classes. Other participants also could not transcend a peripheral procedure for assessments, such as measurement and grading. Such a lack of understanding of assessments led to a complete misunderstanding of performance assessment, which has led to a culture that emphasized measurement-oriented evaluation even after the introduction of performance assessments.

> "I was really dumbfounded when the performance assessment was first introduced. Similarly-nuanced different terms like performance assessment, authentic assessment, or alternative assessment kept pouring into the fields of education. There's not much difference compared to previous PE performance assessments, since these started without accurate instructions on how to conduct the assessment." (Tae-ho)

> "Those modifiers in front of the word 'assessment,' like course-oriented, comprehensive, or consistent, put quite a lot of pressure on teachers. They delineate the concepts but it would take significant efforts to have a clear understanding as to how to assess. Most realistically, teacher training would be a measure for upgrading their expertise, and yet PE training has not been properly conducted. Despite its name being PE assessment training, most of the training was practically not much more than just a one-time event, where ideas for class contents were shared." (Woo-jin)

The interviews with Tae-ho and Woo-jin revealed that the performance assessment system was initiated without an adequate introduction of the concept of performance assessment, and that insufficient training for currently working teachers was provided. Common issues were identified in that not enough opportunities were available for the PE teachers compared to those teaching knowledge-based subjects, and the training programs currently provided are decontextualized, one-time day training courses into which comprehensive instruction-oriented PE contents are poured (Armour & Yelling, 2004). Such problems in teacher training have dissuaded teachers from developing their assessment literacy. Therefore, teachers have conducted "instruction without assessments" and have rated their expertise on

assessment lower than that on instruction. In this regard, Soo-won stated:

> "Students' parents have less interest in PE assessments compared to their interest in knowledge-based subjects. Rather, I think that designing a PE class in a more fun and safe manner would better meet the needs of students and their parents. In this respect, I believe class expertise is more important than assessment expertise. Boring classes or classes where accidents occur result in complaints, but no complaints have been raised related to PE assessments." (Soo-won)

As seen in the above interview, Soo-won regarded instruction and assessments as mutually exclusive. This misconception is prevalent among teachers, and instruction and assessments are mutually interdependent. An assessment is a process that collects information not only about students' learning behaviors but also about teachers' teaching behaviors (McMillan, 2007; McMillan & Workman, 1999). Therefore, a good assessment process can enhance teachers' teaching abilities. Although teachers are only concerned with instruction for their own classes, conducting correct assessment based on their assessment literacy would increase their instruction capabilities. Furthermore, in the context of South Korea, where national-level physical education curriculum is specified in terms of its design, planning and evaluation the interaction between instruction (or pedagogies) and forms of assessments is strong. This is because the fit-for-curriculum assessments that actively incorporate what the national education program intends to achieve require a close-knit alignment between education course, class, and assessment (Penney et al., 2009). Consequently, it would be possible to enhance the efficiency of an education course if teachers' expertise, including assessment literacy, is properly developed.

### Assessment Application: Labor exchange and recycling

Assessment application means that teachers' knowledge about effective assessment is formed in a classroom, through processes that help collect evidence for assessment interpretation (Hay & Penney, 2013). Generally, assessment application begins with formulating an assessment plan. However, research participants revealed a "labor-exchange and recycling" culture in the assessment planning phase, in which teachers share their work with others, as well as recycling that which is used in other classes. Ha-jin and Jong-soo said:

> "Except for two to three main subjects that specialist teachers teach, higher grade lead teachers generally need to come up with their assessment plan for eight to nine subjects. Too many. For this reason, lead teachers who teach the same grade divide their subjects and plan for their assessments. It's sort of like 'two hands are better than one.' Since they are assigned subjects based on their expertise, they are able to plan and implement quality assessments. Wouldn't it be improbable for us to be experts in all subjects?" (Ha-jin)

> "When planning for their performance assessments, many teachers either refer to, or use the entirety of the previous year's assessment plan, or plans shared on teachers' online communities. Good for saving time needed for planning, and reliable since those plans are made by teachers with expertise in PE." (Jong-soo)

As indicated above, it is not necessarily wrong for those with expertise on specific subjects to map out the performance assessment on those subjects and for teachers to share quality assessment plans. As seen in research participants' interviews, this could lead to a self-developing culture arising from a long-standing trial-and-error process led by the communities of elementary school teachers who need to teach various subjects by themselves. Planning a performance assessment should be preceded by a process that considers students' developmental levels, resources available, and consultations with teachers. All the previous performance assessments are the byproducts of such a process, from the previous year, and so some would consider them to be verified assessment plans, even though they do not take into account the current students. Furthermore, assessment plans shared through online teachers' communities often elicit replies from teachers who have actually implemented those plans, thus facilitating reduction of trial and error based on the simulations of the implementation process and results. These qualities will help utilize assessment plans in a labor-exchange and recycling manner.

However, such a labor-exchange and recycling style assessment culture has several issues. First, class assessments should be conducted in a class-specific manner. As for classroom assessments, tasks based on the contexts of students' actual life should be assigned (MacMillan, 2007; Baron, 1995), and for this to happen, aspects such as learners' starting behaviors, classroom ecological aspects, and available resources must be considered. A performance assessment planned in a labor-exchange fashion has issues that are more likely to induce identical assessment contents, methods, and procedures from all classes of each grade. A recycling style assessment also ignores various internal and external circumstances of classes, such as the revision of educational processes and changes in learners' demands.

Second, such a labor-exchange and recycling style assessment culture indicates someone else's assessment literacy, not that particular teacher's. With more teachers majoring in Sports Pedagogy, the quality of teacher guide books and other learning material has improved. In addition, the culture of online teachers' communities, where they share feedback on materials has assumed the form of peer scholarship. Nevertheless, it is doubtful whether sharing these assessment materials would be conducive to teachers' developing their own assessment literacy. It is important to identify whether a direct and holistic coverage of all the assessment procedures, ranging from planning to self-reflection relating to assessments, or indirect and partial involvement through sharing materials, would be more effective for developing teachers' assessment literacy.

### Assessment Interpretation: A lack of feedback

Assessment interpretation is the process of interpreting collected information during the application of the assessment. In this procedure, it is necessary for teachers

to be equipped with assessment literacy, to determine their students' levels, and to chart directions for their future engagement and a groundwork for appropriate curricular and pedagogical adjustments. The results of assessment interpretation are transmitted to the students as feedback, so it is very important that the teacher's feedback includes these cues and that it is provided at the appropriate time.

Data analysis reveals that there has been a pervasive tradition among research participants to conduct an assessment during the last period of a unit and examine their assessment results at the end of the semester, which does not give them time to give feedback to the students. In this regard, research participants Shin-young and Tae-ho reported as follows:

> "If 12 T-ball classes are scheduled, a performance assessment will be conducted in the final (12th) class. This is because assessing during the last class would allow me to accurately assess which stage a student has reached." (Shin-young)

> "I usually leave the examination of assessment results to the semester's end. Based on my grading students on the day of the performance assessments, their final grades are decided and texts composed of four or five sentences are produced. On the last day of the semester, students and their parents are notified." (Tae-ho)

The analysis on other research participants' plans regarding performance assessments showed little difference in that they also measured assessment contents during the last unit, though with a bit of a time difference, and that they wrote and notified the assessment results at the semester's end. This signifies that our education fields have not been able to transcend Tyler's viewpoints-based assessment method, which regards the education process, teaching-learning, and evaluation cycle from a linear and sequential perspective. The most serious problem in Tyler's style of the output-oriented summative assessment lies in focusing on the core objective of assessments, that is, "feedback." Conducting the evaluation at the end of the semester means that the feedback on evaluation results would be generated after the completion of the course, due to which the students would not have the opportunity to amalgamate feedback into their performance. Feedback would be most effective when presented on time (Capel & Whitehead, 2015). Timely and appropriate feedback enhances students' performance ability and contributes to teachers' deciding and making arrangements about students' future assignments, their teaching contents, and methodology. When the formative assessment perspectives from which teachers record students' performance and give feedback to their students are maintained at each unit, timely and appropriate feedback can be formed (Earl, 2003). An assessment can be influential when a teacher becomes an effective mentor, kindred guide, and accurate reporter (Wilson, 1996).

*Critical Engagement with Assessment: Teachers' critical thinking without pedagogy*

The main points of teachers' literacy in relation to critical PE assessments are to recognize the importance and repercussions of the assessments, and to challenge the naturalness of assessment practice, performances, and outcomes. Therefore, they must challenge existing assessment methods in PE and consider the ramifications of the sociocultural assessments.

The neutrality of assessments has long been a dilemma. Technically, assessments cannot be neutral. Once the decision as to "what to assess" has been made, its neutrality is lost. Thus, teachers should predict which student would benefit and which student would be negatively influenced, and then formulate alternative approaches. Research participants had the following concerns with regard to the neutrality of their assessments:

> "It often happens that students with better physical health exercise get better grades in their performance assessments. Unlike other subjects, PE should be a subject in which students are free from assessment and able to have fun. The last thing I would like to see would be kids with lower kinetic functions having low PE grades. Therefore, I would give my students relatively better grades than in other subjects." (Ha-jin)

> "I don't think it's right for students' socio-economic status to have any influence on PE, or even on knowledge-based subjects. That's why I am trying to give them a better score on assessments if they are not able to go to private sports clubs or have sports experience with their parents, just to be fairer and more considerate with them." (Jae-in)

It cannot be said that research participants' linking the subject characteristics of elementary school PE to being "fun," and their considerations for children of a lower socio-economic status ruin the neutrality of the assessment. This can be considered an awareness of the purpose of PE subjects and critical thinking from a sociocultural perspective. However, such issues are not being considered in a pedagogical manner. Issues on prerequisite learning are also raised, as there are students who have a higher starting level in all subjects. Conducting a "give-away" assessment under the pretexts of distinct characteristics of PE and sociocultural influencing factors may spawn the undesired consequence of dragging the status of PE lower.

More focus should be placed on "process-oriented assessments." Though not a part of assessment results, the focus should be on how much students have grown compared to their starting behaviors in an assessment process (Hortigueala et al., 2016). Nonetheless, research participants are trying to be considerate, measuring results strictly from the perspective of assessment of learning. A prerequisite for the genuine process-oriented assessment is assessment for learning. This means being pedagogically considerate from an assessment-for-learning perspective to prepare a class strategy for students who have no prior experience; this would help them participate in assessments of their enhanced performance.

Furthermore, based on their critical thinking about existing PE and performance assessments, research participants endeavored to remove themselves from the framework. What they commonly selected for alternative assessments were self-assessments and peer-assessments, that is, student-led assessments. On the one hand, self-assessments would allow for critical

thinking and self-reflective awareness about their learning (Gordon, 1992; Hill & Ruptic, 1994), and peer-assessments could enhance the connectivity between the assessment process and learning (Green, 1994), both of which are somewhat encouraging. On the other hand, research participants' self- and peer assessments have critical issues. Comparison of self-assessments and peer-assessments records and the final performance assessment outcome document indicated that the research participants did not incorporate students' self-assessments and peer-assessments into the final document. Min-jeong and Soo-won stated the following reasons:

> "I sometimes conduct self- and peer assessments because they have their own share of effectiveness, but it's hard to put my trust in what my students have assessed themselves. In class, was the person my teammate, and out of class, was the person my friend? These factors have a more significant influence in PE than in any other subject." (Min-jeong)

> "The issue of neutrality doesn't only happen in peer assessments. There's this case where students who are favorable to other students' assessments apply strict and harsh standards to their own assessments." (Soo-won)

Several studies raised concerns regarding students' PE assessments, particularly on issues of neutrality (Brennan, 2001; Freeman, 1995; Porter & Cleland, 1995). However, given the fact that research participants decided on their students' final grades with a consideration of students' physical abilities and socioeconomic status when conducting assessments on affective fields, and that they also favored those students who had actively participated in class, who helped prepare the class, and who had behaved well (Black & Dockrell, 1980; Frey & Schmitt, 2010), the issue of outwardly appearing neutrality plagues both teachers and students alike. In addition, the discrepancy between the reliability of teacher assessment and of student assessment, to which research participants are referring, is also pertinent to consider. Probably, this would not involve the difference between teachers and students, but the difference in what each group has understood regarding assessment literacy.

An alternative to remedy teachers' distrust in student-led assessments would be to make their students assessment-literate (Hay & Penney, 2013). For that to occur, teachers must endeavor to improve their students' assessment literacy. It is necessary to provide students with more opportunities for peer assessments and to involve them in the entire assessment process, from communicating the purpose of the assessments to the utilization of results. Porter and Cleland (1995) mentioned that, initially, students become emotional during their peer-assessments, but subsequently, they partake in the essence of assessments. In addition, Freeman (1995) and Brennan (2001) pointed out the importance of assessor training to increase the credibility of student-led assessments. Students' assessment literacy can be enhanced only by more frequent participation in assessments.

## Conclusions and Suggestions

This study examines the issues related to PE assessments in elementary schools from an assessment literacy perspective. Four main issues rose to the surface (*a lack of understanding of performance assessment, recycling of previous assessments, a lack of feedback to the students, and Teachers' critical thinking without pedagogy*); this information adds evidence to support the need to develop assessment literacy of both teachers and students. Based on this information, I would like to present three ways to address these issues:

First, the educational paths for current teachers to develop elementary school teachers' assessment literacy should be diversified. PE training programs for those currently teaching in the field for elementary schools are quantitatively and qualitatively insufficient (Lawrence, 2003; Harris et al., 2012). In-service teacher education, geared towards enhancing expertise in assessment, has been scarce (Stinggins & Conklin, 1992; Nitko & Brookhart, 2011). Developing assessment literacy should not be solely dependent on a teacher's individual efforts. It is necessary to provide in-service teacher education that can share assessment-related professional techniques and provide consistent learning opportunities in the contexts of both schools and educational communities (O'Sullivan & Deglau, 2006; Pritchard & Marshall, 2002).

Second, students should be provided with a specific guideline that leads them to become active evaluators. Teachers should not be satisfied with students in PE classes who are simply "busy, happy and good" (Placek, 1983). When a student becomes an active assessor, the assessment itself can impart learning (Black & William, 2006; Earl, 2003). Thus, a guideline that can enhance students' assessment literacy through theoretical research and on-site studies should be implemented in the relevant fields.

Finally, it is essential to assess the assessment process. The research participants had an opportunity for self-reflection as managerial level teachers or managers in the phases where they were planning and handling results. However, to increase the quality of performance assessments, the entire assessment process should be evaluated from the perspective of being tailored for the curriculum, class, and instruction. Assessment of the assessment process would strengthen teachers' literacy and equip them with a self-reflective opportunity to determine what has been missing in their literacy.

## References

Annerstedt, C. & Larsson, S. (2010). 'I have my own picture of what the demands are...': Grading in Swedish-problems of validity, comparability and fairness. *European Physical Education Review*, *16*(2), 97-115. doi:10.1177/1356336x10381299

Armour, K. M. & Yelling, M. R. (2004). Continuing professional development for experienced physical education teacher: towards effective provision. *Sport, Education and Society*, *9*(1), 95-114. doi:10.1080/1357332042000175836

Baron, M. (1995). *Authentic assessment: The key to unlocking student success*. Lancaster, PA: Techonomic.

Black, H. D. & Dockrell, W. B. (1980). Assessment in the Affective Domain. *British Journal of Educational Research*, *6*(2), 197-208. doi:10.1080/0141192800060208

Black, P. & William, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81-100). London: Sage.

Brennan, R. L. (2001). Generalizability theory. New York: Springer.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximise student learning*. Thousand Oaks, CA: Corwin.

Fullan, M. (2002). The role of leadership in the promotion of knowledge management in schools. *Teacher and Teaching: Theory and Practice*, *8*(3), 409-419. doi:10.1080/135406002100000530

Freeman, M. (1995). Peer assessment by group of groups work. *Assessment & Evaluation in Higher Education*, *20*, 289-300. doi:10.1080/0260293950200305

Frey, B. B. & Schmitt, V. L. (2010). Teachers' Classroom Assessment Practices. *Middle Grades Research Journal*, *5*(3), 107-117.

Gordon, M. J. (1992). Self-assessment programs and their implications for health professions training. *Academic Medicine*, *67*, 672–679. doi:10.1097/00001888-199210000-00012

Green, M. (1994). Children as evaluation understanding from the inside. In Cambourne, B. & Turbil, J. (Eds.). *Responsive Evaluation. Portsmouth* (pp. 83-103). NH : Heinemann.

Hardman, K. & Marshall, J. J. (2000). *World-wide survey of the state and status of school physical education*. Manchester: University of Manchester press.

Harris, J., Cale, L., & Musson, H. (2012). The predicament of primary physical education: a consequence of 'insufficient' ITT and 'ineffective' CPD? *Physical Education and Sport Pedagogy*, *17*(4), 367-381. doi:10.1080/17408989.2011.582489

Hay, P. & Penney, D. (2009). Proposing conditions for assessment efficacy in physical education. *European Physical Education Review*, *15*(3), 389-405. doi:10.1177/1356336x09364294

Hay, P. & Penney, D. (2013). *Assessment in physical education: A sociocultural perspective*. London: Routledge.

Hill, B.C., & Ruptic, C. (1994). *Practical aspects of authentic assessment: Putting the pieces together*. Norwood, MA: Christopher-Gordon.

Hortigueala, D., Fernandez-Rio, J., & Perez-Pueyo, A. (2016). Long-term effects of the pedagogical approach on the perceptions of physical education by students and teachers. *Journal of Physical Education and Sport*, *16*(4), 1326-1333. doi:10.7752/jpes.2016.04210

Lawrence, J. (2003). The impact and use of the TOPs programme in one local education authority. *British Journal of Teaching Physical Education*, *34*, 44-48.

Matanin, M. & Tennehill, D. (1994). Assessment and grading in physical education. *Journal of Teaching in Physical Education*, *13*(4), 395-405. doi:10.1123/jtpe.13.4.395

McMillan, J. H. (2007). *Classroom assessment: Principles and practice for effective standards-based instruction* (4th ed.). Boston: Allyn & Bacon.

McMillan, J. H. & Workman, D. (1999). *Teachers' classroom assessment and grading practices: Phase 1*. Richmond, VA: Metropolitan Educational Consortium.

Nitko, A. J. & Brookhart, S. M. (2011). *Educational assessment of students* (6th ed.). MA: Pearson.

O'Sullivan, M. & Deglau, D. (2006). Principles of professional development. *Journal of Teaching in Physical Education*, *25*, 441-449. doi:10.1123/jtpe.25.4.441

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Penney, D., Brooker, R. Hay p., & Gillespie L. (2009). Curriculum, pedagogy and assessment: three message systems of schooling and dimensions of quality physical education. *Sport, Education and Society*, *14*(4), 421-442. doi:10.1080/13573320903217125

Placek, J. H. (1983). Conceptions of success in teaching: busy, happy, and good? In Templin, J & Olsen, J. (Eds.). *Teaching in Physical Education* (pp.45-56). Champaign, IL: Human Kinetics.

Porter, C. & Cleland, J. (1995). *The portfolio as a learning strategy*. Portsmouth, NH: Boynton/Cook.

Pritchard, R. J. & Marshall, J. C. (2002). Professional development in 'healthy' vs. 'unhealthy' districts: top ten characteristics based on research. *School Leadership and Management*, *22*, 113-141.

Stiggins, R. J. & Conklin, N. F. (1992). *In teachers` bands: Investigating the practices of classroom assessment*. Albany, NY: SUNY.

Thompson, M. D., & Penney, D. (2015). Assessment literacy in primary physical education. *European Physical Education Review*, *21*(4), 485-503. doi:10.1177/1356336x15584087

Wilson, R. (1996). Assessing students in classrooms and schools. Toronto: Allyn & Bacon.

Yoo, J. (2005). Understanding the developing process of physical education curriculum at the secondary physical educator level. *Korean Journal of Sport Pedagogy*, *12*(1), 31-56.

This page is intentionally left blank

www.**iejee**.com