HOSTED BY
ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Journal of Sport and Health Science 9 (2020) 149–151
www.jshs.org.cn

Commentary

A clash of fundamental assumptions: Can/should we measure physical literacy?

Ang Chen

Department of Kinesiology, University of North Carolina at Greensboro, Greensboro, NC 27412, USA

Received 7 September 2019; revised 20 October 2019; accepted 18 November 2019
Available online 21 November 2019

The special issue of the *Journal of Teaching in Physical Education* (*JTPE*)[1] on physical literacy (PL) has provided much needed information about measuring and assessing student learning in terms of the well-defined and yet abstract concept. Multiple authors have elaborated from diverse educational and psychometric perspectives on the development of the construct, the measurement model, and the assessment of the construct. In this attempt, many issues associated with the conception of PL were clarified and insightful information and arguments were brought forward. In this regard, the special issue has moved our understanding of the concept one step further. More importantly, it points to a possibility to actualize the teaching of the concept in physical education (PE).

A core collection of articles is focused on specific issues related to assessing PL in PE. Cairney et al.[2] presented a case of validating a PL construct with a sample of 5th- and 7th-grade students in Canada. Shortt et al.[3] and Keegan et al.[4] conducted separate studies in the United States and Australia on defining PL domain specificity and content representativeness using Delphi or modified Delphi methods. Barnett et al.[5] developed and articulated the guidelines for selecting measures from existing assessments in PE to assess PL. Other articles also touched on PL assessment to imply that PL should and can be assessed in PE. Although the effort is plausible, establishing valid PL assessment tools and systems may be a much more complicated endeavor than what is being attempted and presented in the special issue. The endeavor requires careful considerations that seem absent in the validation studies. The purpose of this commentary is to raise a few important concerns about the current practices of creating PL assessment, and challenge the use of the traditional measurement theory in developing assessment tools to measure PL as manifested in the special issue. I believe that an in-depth discussion will help us to clarify these concerns about PL measurement and inform future efforts in developing assessment systems for this very difficult-to-measure concept.

In the following, I will first briefly share my understanding of the PL concept, then elaborate on its philosophical root of the monism as the fundamental assumption. After that, I will discuss the conceptual conflict between the PL assumption and the fundamental assumption of the measurement theory. Further, I will point out why we are deeply challenged by the conflict and the dilemma we are facing in terms of assessing PL in PE. Lastly, I will try to articulate where we should go with PL assessment in PE. I hope that this commentary will lead to a productive discussion that will result in a consensus as well as a platform upon which PL can be carefully conceptualized for students to experience in PE and potentially operationalized for educators to assess.

1. My understanding of PL and its assumption

PL seems to have many forms of definitions and interpretations in the literature. Several authors in the *JTPE* special issue emphasized that PL has been around for a long time and, during this time, has been taking quite a few forms of conceptions.[6] Although this fact is true, it is unarguably apparent that PL has never received much scholarly attention in terms of assessment until Margaret Whitehead[7] proposed the current PL concept. A careful read of the articles about PL assessment in the special issue (and other articles as well) will lead one to the same conclusion: All assessment efforts described in the special issue seem to be guided by the PL conception proposed by Whitehead.[7] This fact intrigued my thinking about the issues I am presenting below and delimits my discussion within the conception of the current PL.[7]

When defining PL, Whitehead[7] clearly stands on a holistic conceptualization: "physical literacy can be described as the motivation, confidence, physical competence, knowledge, and understanding to maintain physical activity throughout the life course" (p. 11–12). She further clarifies that the foundation for PL is monism that should be understood as "different modes of our 'body' or embodiment and the interrelationships between dimensions" (p. 19) which she refers to as the above 4 elements. Built on the monism foundation, there are 3 fundamental premises of PL. First, PL is characterized by monism, "Physical literacy can *only* be conceptualized in the context of

Peer review under responsibility of Shanghai University of Sport.
E-mail address: a_chen@uncg.edu

https://doi.org/10.1016/j.jshs.2019.11.002

monist approach to human being. . . *There is no sense* in which we can talk about our embodiment or our 'body' as being a *discrete aspect* of our personhood" (p.19, italic added). Second, *embodiment* takes place in "modes" in the PL journey, which is signified by the idea of "*body as lived*" that is apart from the "*body-as-object*" point of view for human physical movement. She elaborates that the body-as-lived embodiment can be "pre-reflective or pre-conscious", which is "an example of monism in action" (p. 19). Human physical movement should be and must be a holistic experience regardless of our consciousness, "Physical literacy is founded on and relies on an appreciation of embodiment in both (conscious and pre-conscious, author added) modes" (p. 19). Third, the dimensions of PL are dynamically and holistically interrelated. Changes in one "will, in almost every case, have an impact *across all other dimensions*" (p. 19, italic added). This dynamic interrelationships among human capacity goes beyond the dimensions of PL into other human capacities such as cognition, creativity, and more. These premises are articulated further in Chapter 3 of Whitehead's text to strengthen the monism foundation of the PL concept.[7]

2. The fundamental measurement assumption and the conflict

Measurement and assessment take on different epistemological perspectives as opposed to that of PL. Although measurement is concerned with objectively assigning scores to performance so that a judgment can be made with confidence, assessment is concerned with judgmental interpretation of the scores representing the performance. In other words, measurement provides numerical evidence for assessment. At the center is the issue of validity that signifies the extent to which the scores represent the performance. Since its inception, measurement relies on a fundamental assumption that an entity to be measured can be broken down into small and independent measurable pieces as elements or components for objective observation and scoring. Breed,[8] in defending measurement applications in assessing educational achievement, made this assumption clear by stating "Measurement based on the method of analysis is, therefore, said to deal with relatively insignificant *pieces* of personality (italic added), and neglects that latest fetish of the legions of educational light, the integrated individual" (p. 119). He hoped that "educational measurement, by virtue of its analytical power, discovers and thereby contributes to the later improvement of integrations" (p. 120). The integrations he was referring to are the whole, resulting from combining pieces of measurable components. Many sub-assumptions in measurement grew out of this key assumption to ensure that assigned scores or numerical ratings are independently meaningful and useful in reference to the concept/construct for which the scores and ratings are developed. These sub-assumptions include those about content domain specifications, sampling, construct development, scaling (numerical scoring), drawing inferences, and more.[9] In my opinion, this key measurement assumption is in direct epistemological conflict with the monism premise because it is impossible and makes no sense, using Whitehead's language, to break up PL into *independent* components merely for measurement and at the same time maintain the monism premise.

Validity is at the core of the measurement theory. The current measurement practices presented in the special issue point to 2 specific threats to validity. The first threat is a conceptual one, as discussed, related to the opposing views of the monism premise and the fundamental measurement assumption that are contradictory to each other. In this regard, the literature on PL often questions or challenges the assessment practices. For example, Corbin[10] notices the disagreement among scholars about what should be emphasized in PE under the PL framework. He further points out that available PL assessment systems are actually assessing PE content rather than PL. In short, the issues concerning validity include conceptualization, scaling, *etc*. in developing the assessment systems. Another issue related to this threat is the scaling and scoring methods adopted in the assessment systems. Most PL assessment systems use a simple arithmetic additive method to integrate scores from all dimensions to arrive at an aggregated total score to represent PL levels. This method has been challenged and deemed unacceptable in assessing complex performances[9] because of the aforementioned threats to validity. Based on the monism premise and PL's holistic view of human movement experience, mathematically integrating the scores from the dimensions (dimensional scores) cannot best estimate PL in its entirety. In other words, a total score made by aggregating sub-scores may not represent PL. Theoretically speaking, if a construct/concept can be clearly defined and operationalized for measurement, a composite score with appropriate weights assigned to its components may be adequate in representing the construct. In the case of PL, however, such a composite score may not represent the construct due to the indivisible nature of the PL concept and its components based on the monism assumption. If PL ought to be assessed in PE, the assessment system should be aligned, or at least should not be in conflict, with the monism assumption to reflect the holistic "body as lived" embodiment as learning achievement. Thus, the current practice of using a composite score aggregated mathematically from dimensional scores may not be indicative of the fulfillment of this experience.

The second threat to validity is methodological. Procedurally in validation research, after a domain specificity study with experts determining the evidence needed for validity in relation to test content (PL), a known-group method is often the choice of methods to provide validity evidence with the targeted population that the assessment is developed for, in this case K−12 students. The validity evidence should be based on the content as well as relevant to the content (see below for why this is particularly important) and should be supported by a variety of cumulated evidence. At this step of validation, data may be compared between a group of students that we know possess the targeted characteristics, in this case PL, and a group that we know do not possess the characteristics. The comparison will lead to a conclusion about the assessment system's sensitivity of detecting the presence or absence of the characteristics in terms of group membership and/or the sensitivity of determining the levels of the characteristics. A lack of this sensitivity will deem that the assessment system will not be able to provide valid scores for judgment, which is interpreted as lack of validity.

A dilemma here is that based on the definition of PL, few K−12 students can be classified as physically literate with whom the known-group method can be applied for this validation purpose. Corbin[10] raised this concern by questioning Canada's goal for every child to become physically literate by the age of 12. The reasoning is that if this goal has not been achieved, validation studies with this population, as reported in the special issue,[2] will not be able to provide trustworthy validity evidence for the assessment system. Using such an assessment system, we will not know for sure whether a student is physically literate! Although the results[2] indicate that the measurement model is fit with acceptable goodness-of-fit indices, the issues of being fit for what and whom (validity) remain. Until the assumption conflict is resolved and the dilemma is settled, the validity of the assessment system remains in question. This reasoning seems to lead to a conclusion that the model may not be useful for assessing 5th- and 7th-grade students because the sample is unlikely from a population that we have confidence to declare to be physically literate.

One observation I need to reiterate is that the above discussion is based on an understanding that the PL concept presented in the special issue is indeed based on Whitehead's conceptualization.[7] Almost all authors stated that the PL concept which they relied their work on may or may not be consistent with Whitehead's original idea. Indeed, Cariney et al.[6] state that the concept in the special issue can be conceptualized as an evolving "trans-disciplinary" construct of PL (p. 82). What is not clear is which components or sub-components in the PL concept have been or are being transformed and what outcomes resulted from the transformation other than the Whiteheadian conception that is now being acknowledged and adopted widely in the literature including the special issue. In the PL assessment articles in the special issue, the sub-concept names, terminologies, and definitions are almost identical to those by Whitehead.[7]

3. What is possible moving forward?

Lundvall[11] has warned that moving forward with PL as a framework for PE is a road never traveled and full of challenges. On the other hand, she encourages scholars and practitioners to embrace the idea because the concept presents a tremendous possibility to change the prospect of future PE. This commentary presents one of the challenges that we are facing.

How should we address the challenge? Where should we go from here? I believe that the power of PL lies in its guiding purpose that is philosophical and forward thinking. Metaphorically, PL presents a journey of life-long PE that individual learners travel on but may not be able to arrive at its destination while in school. Can and should we accept this metaphor? In my opinion, we have few choices but to accept it. We may not be able to, and perhaps should not, operationalize the PL concept, as articulated so well as a holistic entity, merely for measurement or assessment purposes. In other words, if Whitehead's PL conceptualization is honored, assessment of it may not be attempted. The current assessment attempts have not resolved the conflict between the fundamental assumptions between monism and measurement. Thus, defending these assessment attempts can be difficult. In other words, we may fail to counter the argument that these assessments may be assessing the pieces (knowledge, skill, and fitness, *etc.*) in the conventional PE curriculum rather than assessing PL as conceptualized and articulated by Whitehead.

Shortly after Society of Health and Physical Educators (SHAPE-America) adopted PL as an overarching goal for PE, a group of scholars discussed the possibility to operationalize it for PE in a special issue of the *Journal of Sport and Health Science*.[12] The authors expressed diverse perspectives that embraced, challenged, and questioned the possibility to integrate the PL concept into K−12 PE. Nevertheless, the implications are clear from that discussion; that is, the PE curriculum might be the place to start an operationalization journey.

I consider PL as an ever-lasting, futuristic concept. It will remain futuristic because it describes a journey rather than a destiny. It is a completely different conceptualization for physical movement to be lived by all human beings. It appears that if we want to use it as a guiding framework and a goal for PE, we may need to drastically change the curriculum first to adapt to the idea of monism, help students holistically learn/experience movement knowledge, motivation, fitness, skills, and values as a monist experience. Only after a new curriculum is conceptualized, an assessment system be developed as part of the curriculum development. Even in this case, one should realize that the assessment system is to assess students' learning achievement of curriculum goals rather than assess PL itself, because only the moving individual can tell whether she/he has arrived at the PL destination.

Competing interests

The author declares that he has no competing interests.

References

1. Dudley D, Cairney J, Goodway J. Special issue on physical literacy: evidence and intervention. *J Teaching Phys Educ* 2019;**38**:77–8.
2. Cairney J, Clark H, Dudley D, Kriellaars D. Physical literacy in children and youth − a construct validation study. *J Teaching Phys Educ* 2019;**38**:84–90.
3. Shortt CA, Webster CA, Keegan RJ, Egan CA, Brian AS. Operationally conceptualizing physical literacy: results of a Delphi study. *J Teaching Phys Educ* 2019;**38**:91–104.
4. Keegan RJ, Barnett LM, Dudley DA, Telford RD, Lubans DR, Bryant AS, et al. Defining physical literacy for application in Australia: a modified Delphi method. *J Teaching Phys Educ* 2019;**38**:105–18.
5. Barnett LM, Dudley DA, Telford RD, Lubans DR, Bryant AS, Roberts WM, et al. Guidelines for the selection of physical literacy measures in physical education in Australia. *J Teaching Phys Educ* 2019;**38**:119–25.
6. Cariney J, Kiez T, Roetert EP, Kriellaars DA. 20th-century narrative on the origins of the physical literacy construct. *J Teaching Phys Educ* 2019;**38**:79–83.
7. Whitehead M. *Physical literacy throughout the lifecourse*. New York, NY: Routledge; 2010.p.3–29.
8. Breed FS. Fundamental assumptions in educational measurement. *Phi Delta Kappan* 1937;**20**:118–23.
9. Delandshere G, Petrosky AR. Assessment of complex performances: limitations of key measurement assumptions. *Educ Res* 1998;**27**:14–24.
10. Corbin CB. Implications of physical literacy for research and practice: a commentary. *Res Q Exerc Sport* 2016;**87**:14–27.
11. Lundvall S. Physical literacy in the field of physical education − a challenge and a possibility. *J Sport Health Sci* 2015;**4**:113–8.
12. Chen A, Sun H. A great leap of faith: editorial for *JSHS* special issue on physical literacy. *J Sport Health Sci* 2015;**4**:105–7.