



UNIVERSITY OF
GLOUCESTERSHIRE

This is a peer-reviewed, post-print (final draft post-refereeing) version of the following published document, © 2023 SHAPE America This is an Accepted Manuscript of an article published by Taylor & Francis in Research Quarterly for Exercise and Sport on 10.02 2023, available at: <https://doi.org/10.1080/02701367.2022.2125927>. and is licensed under Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0 license:

**Gandrieau, J., Schnitzler, C., Cairney, J., Keegan, R., Roberts, William M ORCID: 0000-0001-5736-5244, Barnett, L.M., Bentsen, P., Dudley, D.A., Raymond Sum, K.W., Venetsanou, F., Button, C., Turcotte, S., Berrigan, F., Cloes, M., Rudd, J.R., Riga, V., Mouton, A., Vašíčková, J., Blanchard, J., Mekkaoui, L., Derigny, T., Franck, N., Repond, R-M., Markovič, M., Scheuer, C. and Potdevin, F. (2023) Development of ELIP to Assess Physical Literacy for Emerging Adults: A Methodological and Epistemological Challenge. Research Quarterly for Exercise and Sport, 94 (4). pp. 1169-1182.
doi:10.1080/02701367.2022.2125927**

© 2023 SHAPE America This is an Accepted Manuscript of an article published by Taylor & Francis in Research Quarterly for Exercise and Sport on 10.02.2023, available at: <https://doi.org/10.1080/02701367.2022.2125927>.

Official URL: <http://doi.org/10.1080/02701367.2022.2125927>

DOI: <http://dx.doi.org/10.1080/02701367.2022.2125927>

EPrint URI: <https://eprints.glos.ac.uk/id/eprint/12306>

Disclaimer

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.

**Development of ELIP to assess physical literacy for emerging adults:
a methodological and epistemological challenge.**

Words: 7452

Abstract

Following increased interest in physical literacy (PL), development of appropriate tools for assessment has become an important next step for its operationalization. To forward the development of such tools, the objective of this study was to build the foundations of the *Évaluation de la Littératie Physique (ELIP)*: designed to help reduce existing tensions in approaches to PL assessment, that may be resulting in a low uptake into applied settings. We followed two steps: (1) the development of the first version of ELIP by deploying a Delphi method ($n=30$); and (2) the modification of items through cognitive interviews with emerging adults ($n=32$). The expert consensus highlighted four dimensions of PL to be assessed – physical; affective; cognitive; and social – with new perspectives, including a preference for broad motor tests over fitness. Results offer new insights into the assessment of emerging adults' PL, but ELIP still requires further work concerning validity, reliability, and sensitivity.

Keywords: Assessment, Cognitive interview, Delphi method, Validation

Recently, the concept of Physical Literacy (PL) (Whitehead, 2001, 2010) has gained increased attention in international literature (Young et al., 2019). For many, PL has become a key concept in the promotion of lifelong physical activity (PA) in different contexts (e.g., health, sport, and physical education). Nevertheless, question of how to assess and measure PL remains controversial (Barnett et al., 2019; Chen, 2020; Edwards et al., 2018; Green et al., 2018; Young et al., 2021) and need to be expand to other audiences ‘across the lifespan’ (Kwan et al., 2019; Liu et al., 2022; Longmuir & Tremblay, 2016). While measures are increasingly being developed and validated for use in childhood assessments (Edwards et al., 2018; Jean de Dieu & Zhou, 2021; Ke et al., 2022), there remains few measures suitable for use in emerging adulthood (Arnett, 2000).

The Physical Literacy: one concept for several definitions.

The International Physical Literacy Association (IPLA) adopts Whitehead’s perspective in defining PL as “the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for engagement in physical activities for life” (IPLA, 2017). Despite its international popularity (Edwards et al., 2017) the following text also considers other perspectives that may provide additional insight into what PL represents (Edwards et al., 2017; Martins et al., 2020; Shearer et al., 2018; Young et al., 2019). In a recent European consensus statement, which included French representatives, PL appeared as “the skills and attributes individuals demonstrate through physical activity and movement across their life course. It can be understood as a process and as an outcome that individuals pursue through an interaction of their physical, emotional, social and cognitive learning” (Physical Literacy for Life, 2021). This definition, more aligned to the Australian conception (Keegan et al., 2019), highlights the necessary holistic development of people through movement, and suggests that different outcomes, or milestones, can be set across the

life course through measurement tools. Around these definitions, two opposing ‘poles’ with clear boundaries can be characterized: an ‘idealistic’ approach - philosophically oriented - and a ‘pragmatic’ one - more focused on bringing the concept into practice (Edwards et al., 2017). Despite these different opinions, recent perspectives emphasize that PL concept “do not need to be ‘set in stone’ or reduced to a ‘single’ definition” but is dependent on the context and the stakeholders for whom it is intended (Young et al., 2019, p. 3). The value of an adaptable and usable concept in a broad range of context such PL surely lies in these multiple approaches (Young et al., 2022).

While trying to preserve the philosophical foundations, we consider PL as a ‘scientific’ (Cairney, 2019a) and pedagogical concept. As most PL experts, our approach rejects a Cartesian distinction between mind and body and rather promote the idea of richly holistic and embodied learning (Pot et al., 2018; Whitehead, 2007). Nevertheless, we highlight the value of milestones setting across the PL journey as suggested by the conception of Keegan et al. (2019) and Physical Literacy for Life (2021) definitions. This stance places our PL conception in the ‘medium-level on the PL ladder of abstraction’ (Young et al., 2019), making it adaptable and usable in different contexts while offering the possibility of retaining the original philosophical underpinnings. In this paper, we claim that a reliable, sensitive, and valid assessment – with moderate classification and framing (Young et al., 2021) - will be a cornerstone of this approach useful to design quality PL interventions and to enhance the scientific proof of concept – which in turn will enrich the practice.

The challenge of broadening Physical Literacy for the emerging adults.

By trying to capture PL through different dimensions, some tools which match with the holistic perspective of PL assessment have been proposed recently – IPLA charting (Whitehead, 2019); PL-C Quest (Barnett et al., 2020); CAEPL (Chen et al., 2020b); PPLA-Q

(Mota et al., 2021); - and are added to well-known tools like the CAPL (Longmuir et al., 2015), PPLI (Sum et al., 2016), PFL (Lodewyk, 2019), PLAY tools (Caldwell et al., 2021), or Pre-PLAY (Cairney et al., 2018). Nonetheless, there remain opportunities to advance measurement capability for PL: both in terms of broadening the population assessed and respecting the philosophical foundations of PL (Pot et al., 2018; Whitehead, 2010).

To date, PL assessment has largely focused on school-age children (e.g., CAPL; PL-C Quest). However, PL is a concept that must be understood throughout the life course (Cairney et al., 2019b; Keegan et al., 2019), and so its assessment also needs to be extended to other age groups according to their characteristics. Emerging adults have been under-represented especially in the PL literature, and they also need to be considered in new measurement tools. In respect to this study, ‘emerging adulthood’ is defined as the time from the end of adolescence to young-adulthood responsibilities, and have to be a target period for PL assessment tools (Edwards et al., 2018; Longmuir & Tremblay, 2016). Many emerging adults leave their family home, begin careers, or enroll in university. It can be considered as a key period during the PL journey, where individuals are, for the last time, engaged in compulsory physical education and where the most active people are still engaged in organized PA (Muller, 2018). The specific characteristics of this period has long been underemphasized. However, it is a crucial and unique period, including changes in brain and body maturation, social skills, friendships, family-oriented socialization, learning about intimacy, and mutual support (Berndt & Savin-Williams, 1993; Hochberg & Konner, 2020). Without appropriate tailored ways of measuring PL in this life-phase, we may fail to understand and support this important life transition for specific population (Arnett, 2000). Literature regarding the PL of this target population is emerging (Kwan et al., 2019), but to date specific assessment instruments are missing, and the other potentially usable tools do not naturally transfer to appropriately assess PL in emerging adulthood. Some pre-existing tools were designed for a

wide age range (e.g., IPLA charting, PLAY tools), but the way of conceptualizing PL for emerging adulthood could be different to children or other age groups (Alipour Anbarani et al., 2022). The tool's language, dimensions, domains, and complexity of items/tasks needs to reflect this specificity. Moreover, the emerging adulthood's way to engage in PA is different from a child, adolescents, or senior (Muller, 2018), as well as the PA commitment process (Bauman et al., 2012).

The epistemological challenge of assessing PL.

Partly due to its development from multiple philosophical perspectives – especially monism, existentialism and phenomenology (Pot et al., 2018) – the assessment of PL is particularly contentious (Chen, 2020; Edwards et al., 2018; Green et al., 2018; Longmuir & Tremblay, 2016; Robinson & Randall, 2017; Whitehead, 2010; Young et al., 2021). The overwhelming majority of assessments (Edwards et al., 2018; Jean de Dieu & Zhou, 2021; Ke et al., 2022; Young et al., 2021) are detached from the holistic perspective that PL is intended to promote - frequently focusing only on motor skills - and the way in which an individual profile of PL development is assessed is a controversial issue (Chen, 2020a; Robinson & Randall, 2017). The linear and simple arithmetic approaches, consisting of summative scores by independent dimension (e.g., PPLI or PLAY) or to attribute less importance to one dimension than to others (e.g., CAPL), are questionable in terms of monist viewpoint underpinning PL (Chen, 2020a). The monist perspective, considering the individual as a whole in which each component interacts with the others with equal importance, could be further enhanced. Assessment procedures can be improved by considering the idea of interdependence and a similar significance between each of the PL dimensions. This corresponds to the complex nature of the human behavior and PL concept (Dudley, 2018;

Heino et al., 2021; O’Sullivan et al., 2020) and tools with ‘strong classification and framing’ do not truly capture this complexity (Young et al., 2021).

Likewise, some tools diverged from the existentialist and phenomenological philosophy advocated by Whitehead. For example, by imposing a restrictive sports-oriented vision of PA (e.g., PPLI) or not really challenging the participant to adapt freely to a complex task (e.g., CAPL), these tools may not have garnered the breadth of information necessary to understand the relationship between individuals and their environment (Whitehead, 2001). Also, to capture the most valuable information in this perspective, a mixed tool combining questionnaires and motor tests could be necessary. Despite the highly pragmatic nature of questionnaires (e.g., IPLA charting, PL-C Quest), it could be considered reductive to focus only on questionnaire responses since it would hardly consider the dynamic relationship between the individual, the environment, and the task (O’Sullivan et al., 2020). The phenomenological perspective assumes that each individual and their relationship to the world is unique (Merleau-Ponty, 1968) and is therefore opposed to a normative and comparative assessment. A ‘medium-level conception on the PL ladder of abstraction’ (Young et al., 2019) would privilege an ipsative evaluation (in comparison only with one’s own previous measure); or ‘charting’ (Whitehead, 2019).

Finally, the identification of the constitutive PL components is also controversial (see Robinson & Randall, 2017; Tremblay & Longmuir, 2017) and needs to be further explored; including for emerging adults. Encouragingly, new PL frameworks (defining statements, domains, elements and guidelines for development, Keegan et al., 2019), philosophical perspectives (O’Sullivan et al., 2020; Rudd et al., 2020), and recent studies on the fundamental/foundational skill determinants of a sustainable physical life (e.g., Hulteen et al., 2018) have all advanced the literature while also reframing the debate about how best one may measure PL as well as in populations other than children. It is crucial to question these

recent reflections for the design of PL assessment tool since “PL and its assessment could have a role to play in opening the domains considered important for lifelong and life wide participation” (Young, 2021, p. 175).

A methodological challenge to produce a valid, reliable, sensitive, and practical tool.

In addition to being a philosophical challenge, developing a new measurement tool for PL presents scientific and methodological challenge (Mokkink et al., 2010). It is widely recognized that the effectiveness of a tool is measured through statistical methods, including focus in terms of validity, sensitivity and reliability (Gunnell et al., 2014; Mokkink et al., 2010). The American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME, 2014, p.11) claimed that validity is “the degree to which evidence and theory support the interpretations of test scores for proposed uses”. The quality of a tool is also verified by its reliability and sensitivity (Mokkink et al., 2010). Validation practices by researchers in sport sciences and sport psychology have been variously critiqued before, challenging the robustness of approaches in this domain (Gunnell et al., 2014; Zhu, 2012). Importantly, achieving the challenge of building an appropriate PL measurement tool would help to promote the concept in the scientific and professional community. Then, a development of a new PL measurement tool requires a rigorous step-by-step approach that needs to be conducted carefully and systematically. By the way, we advocate that the success in the epistemological challenge is largely related to the success in overcoming methodological and statistical challenge.

Hence, to contribute to the ongoing development of PL assessments, we set out to develop a new tool that: (1) captures main philosophical underpinnings of PL; (2) sets milestones for testing PL for emerging adults; (3) enables the combination of subjective and objective assessments; and (4) respects the main steps of a robust validation of a measurement

tool. The insights offered by such a tool will help to design educational programs and to identify learners' development opportunities in order to foster individuals' PL development, while being a positive step in PL journey for each participant. The design of a such measurement tool will also support the theoretical foundation of PL, including identifying whether certain typical PL profiles are favorable to a healthy commitment in PA through emerging adulthood. Thus, the aim of this study was to design the foundation and assess the first validation levels of a PL assessment tool emerging adults (17-21 years old). This study was conducted for the European French context.

Method

We conducted two studies, consecutively, to develop a tool to assess the level of PL development in emerging adults (Figure 1). The first study deployed a Delphi method process (Powell, 2003) to reach a consensus about the structure and the design of the assessment tool (i.e., face and content validity). The second study used a cognitive interview process to ascertain the feasibility and adequacy of the questions' meaning and their interpretation by the respondents (Willis, 2005). This additional insight is rarely included in the development of such assessments, particularly in physical activity and sport sciences, but is necessary to increase reliability and validity (Dietrich & Ehrlenspiel, 2010). All the steps were in consistence with the COSMIN recommendations (Mokkink et al., 2010).

[Figure 1 near here].

Study 1

Participants - Recruitment of expert panel. Our selection process was informed by considerations for (1) recruiting experts, who recently published articles related to PL in journals indexed in PubMed database; (2) recruiting experts of different nationalities; and (3)

gathering different points of view through expertise from different scientific fields. Among 31 experts contacted, 20 responded favorably (66.7% positive responses). We used snowball sampling by inviting experts to propose other experts (Hanson et al., 2020). This resulted in the identification of 15 other individuals not originally identified in the database search. The analysis of their professional background made it possible to include them in the panel of experts and 10 responded favorably (66.7% positive responses). Professional backgrounds included: (1) published papers on PL; (2) published papers on engagement in PA; and (3) involvement in research or professional group on these subjects.

A five-point scale (1= beginner / 5= expert) was completed by each expert to self-assess their PL expertise; the mean response was 4.63 ± 0.49 . The summary description of the participants is shown in Table 1. To thoroughly involve the experts, we offered panel members the opportunity to become co-authors on the final publication generated by the study and included verification-questions in the questionnaires (see section below). In the following reporting and during the process, all data have been anonymized to minimize bias.

[Table 1 near here].

Delphi process to develop ELIP structure. According to Chen (2020a) and Edwards et al. (2017), the debate on how to assess the level of PL needs to be raised among all experts. The goal of the Delphi process, therefore, was used to reach a consensus on a specific research issue after successive rounds of discussions. It is defined as “*an iterative process, designed to combine expert opinion, in order to arrive at a group consensus*” (Keegan et al., 2019, p.3) and is recommended when discordance surrounds a topic (Powell, 2003).

Delphi method's design. We implemented four iterative rounds of Delphi method through the LimeSurvey platform (<https://www.limesurvey.org/fr/>) from May to October 2020. The successive rounds aimed to: (1) examine the constitutive dimensions of the tool; (2) examine the constitutive domains of each dimension; (3) examine tests to assess the

accepted domains; and (4) design a filter to transform the items to be in line with the epistemological foundations of the PL concept. In each Delphi round, the results of the previous round were presented. Experts not responding despite two reminders were removed from the concerned round but were invited to participate in the following rounds ($n < 8$). All the selected participants were considered as PL experts and anchored in different PL perspectives. Thus, no definition was imposed on the panel to design the ELIP, and therefore the result is based on the elements that reached consensus. The experts were informed that the questions asked should be contextualized for the French European context.

Questionnaire design. The literature was screened to identify the potentially relevant dimensions to be included in the ELIP. The first round involved capturing the relevance and comments about the first selection of PL dimensions (Supplementary File 1). The second round consisted of capturing the relevance and comments about a first selection of PL domains for each accepted dimension (Supplementary File 2). Domains are understood as the constituent elements of the dimensions. Based on the results of the previous two steps, the literature was analyzed to identify and/or design relevant tests to assess each accepted domain. For this selection, the research team questioned the content and face validity and examined the relevance and sensitivity for a French emerging adult population of each different tool identified. The relevance of these tests was verified in round 3 (Supplementary File 3). To transform these tests as close as possible to the PL concept, we collected the comments from Round 3 and reformulated them to create ‘filters’ (i.e., sentences that we used to transform the test items). In a last round, experts were questioned about the relevance of the filters (e.g., “*How much do you agree with this idea? The items of social dimension should be causally related to physical activity and not remain general*”). Finally, each accepted item of the selected tests was transformed through the filters until a consensus within the initial team

(GJ, DT, PF, SC) was reached. At the end of this step, we had a set of items to submit to the PL experts for content validation.

For each round, experts responded to an online questionnaire and were invited to rate the relevance of the different propositions thanks to a Likert Scale from 1 to 5 (“*not relevant at all*” to “*very relevant*”). The use of open-ended questions helped to obtain richer data (Powell, 2003), and experts were free to comment on the different domains proposed and to make suggestions.

Consensus requirements. Each round ended only when a consensus was reached. Usually, an item is retained in Delphi method if it is accepted by most of the group’s experts: between 55% and 100% (Powell, 2003). Here, a limit of 66% was tolerated (Powell, 2003) considering the divergence existing in terms of definitions (Edwards et al., 2017; Shearer et al., 2018), as well as the sensitivity of the issue of PL assessment (Chen, 2020a; Edwards et al., 2018). Otherwise, the item was either requested in sub-round if agreement between 61% and 65% or discarded if less than 61%. The comments submitted by the experts were collected and those with at least two converging comments were considered for a next step. Two comments converge when they propose, in a different way or not, a common element (e.g., “*Beliefs should not be in this dimension but in the cognitive dimension*” and “*beliefs are primarily cognitive*”).

Content Validity. Each item reformulated through the filters and instruction was then subjected to content validation by the same 30 experts. Content validity is “*the degree to which the elements of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose*” (Haynes et al., 1995, p. 14). Experts’ judgements on the relevance, representativeness, and clarity were captured through a 1 to 5 Likert scale. We followed the recommendations of Lawshe (1975) to consider the acceptance standard of Content Validity Ratio (CVR) ($n \leq 25$, $CVR \geq 0.37$). At the end of this step, we

had a set of items to be specified by the Cognitive Interview (CI) method with emerging adults.

Results of Study 1

Delphi Method. In total, we invited 46 experts to participate, 31 from literature reviews and 15 via snowball sampling process. We received 30 positive responses (65.2%). Our results are well within the minimum number of 15 experts (Mokkink et al., 2010).

Round 1. The results of the first two rounds are presented in Table 2. The 23 experts (76.6%) who completed the first round strongly accepted to integrate in the ELIP the three following domains: physical (100%), affective (100%), cognitive (95.6%) and slightly less the social dimension (86.9%). According to the experts, all four dimensions must be integrated into the ELIP. Within this study, experts privilege objective rather than subjective tests for the physical dimension and any proposals for other dimensions emerged from the comments left by the experts. Despite the choice to focus on objective tests, experts chose to evaluate perceived aquatic competencies rather than the objective ones (73.9%).

Round 2. The questionnaire was completed by 23 experts (76.6%). In the first sub-round, one affective domain, four social domains and three cognitive domains were directly discarded (agreement<61%). Beliefs (affective), cardiorespiratory capacity (physical) and strength (physical) were re-questioned in a second sub-round (61%>agreement<66%). In this second sub-round, physical fitness (including strength and cardiorespiratory capacity) was discarded, and the belief domain was moved onto the cognitive dimension (Table 2).

[Table 2 near here].

Round 3. The results of the third round are presented in Table 3. Twenty-four experts (80%) completed this round. Fourteen tests were submitted to the experts for their opinion. Eight tests were directly accepted (agreement>66%) and one was directly discarded

(agreement < 61%). Three tests were questioned in a second sub-round according to the comments from the experts and were accepted with slight modifications (e.g., *focusing only on one part of the questionnaire*).

[Table 3 near here].

Round 4. The results of the fourth round are presented in Table 4. We reworded all comments that appeared at least twice in Round 3 as filters. These filters transformed the accepted test items to fit the epistemological foundation of PL. Sixteen (16) filters were drafted and proposed for experts' validation (Table 4). Twenty-five (25) experts (83.3%) accepted thirteen (13) filters (agreement > 66%). None were challenged for a second sub-round.

[Table 4 near here].

Content validity. Twenty-five experts (80.6%) accepted 88 items ($CVR \geq 0.37$). In addition, the completeness of the tool's dimensions and the relevance and clarity of the instructions were also validated ($CVR \geq 0.37$). At this stage, a tool to assess four dimensions (i.e., cognitive, affective, physical, and social) of PL was designed. The first version of the ELIP was constituted with 88 items and three motor tests (Figure 2).

[Figure 2 near here].

Study 2

Cognitive interview procedure. Item readability and comprehension are essential to accuracy in reporting and therefore critical to ensure valid and reliable responses as communication failures are commonplace in questionnaires (Beatty & Willis, 2007). Cognitive interview is a powerful method to understand the thought process used by the identified user-group when answering items, and allow the ability to avoid ambiguity, misunderstandings, and identify unfamiliar wording. It helps to ensure that the ELIP's items are clearly understood by the target participants (Beatty & Willis, 2007) and to verify content

validity of each question through the emerging adults' perspective. In the present study, cognitive interviews were conducted in the French language and context.

Participants. The cognitive interview procedure was conducted with 32 participants (Table 5). They were enrolled and randomly sampled from three specific populations recruited in three high schools and one university in France. This method allows these three populations in consideration with varying academic profiles to be tested: (1) first-year sports science students; (2) students enrolled in traditional program; (3) students enrolled in vocational program.

[Table 5 near here].

Data collection. The interviews were conducted individually by a single interviewer (GJ), who was familiar with the study. The interviewer attended training with an expert interviewer and had to conduct four mock interviews before interviews with study participants began. To achieve a high level of attention from the participants, the cognitive interview was carried out for only one questionnaire per student (from 30 to 45 minutes). According to the consent of the respondents and/or parents, the 32 interviews were audio recorded and transcribed verbatim. We followed Willis' (2005) recommendations to design interview procedures with two steps: (1) introduction and warm-up, and (2) think-aloud and probing techniques. In a first step, a warm-up is conducted to prepare the interviewee to perform a think-aloud interview: *“Try to visualize the place where you live and think about how many windows there are in that place. Can you give me the number of windows please?”*. This question prepares interviewee to the think-aloud method and it illustrates that a poorly formulated or overly general item can be problematic. The interviewer insists on the fact that no perfect answers are expected, but rather that the interview is about testing a questionnaire in progress that includes questions that may be difficult to understand or answer. Participants are reassured by the interviewer's announcement: *“I didn't write these questions, so don't*

worry, I won't be upset if you criticize this questionnaire. All your comments, whatever they are, can help me".

In a second step, the interviews were conducted through think-aloud method mixed with probing method. The think-aloud method asks respondents to verbalize their thoughts and understanding aloud by reading each item and trying to answer the question as they understand it (Supplementary file 4). Reactive verbal probes were used to question participants in response to their behavior (e.g., "*On this question it took you a long time to answer, what happened when you read the item?*"). Pre-planned probes were used to encourage participants to talk aloud about how they understood, processed, and responded to specific items (e.g., "*Can you understand this word?*"). Moreover, retrospective questions were used to identify different elements that are difficult or blurred (e.g., "*Now that you have read the questionnaire, which items were most difficult for you or for one of your friends to complete?*"). In case of misunderstanding, the interviewer tried to get the participant to rephrase the item with their own words. Similarly, the probes were also directed to ensure that the questions asked what they were supposed to measure (e.g., "*Do you think some students will respond 'I' here? What would be the difference between you and them?*"). This procedure was originally designed in three stages but free to continue until reaching theoretical saturation (i.e., sustaining the process until no new findings emerge). Researchers decide by consensus when this saturation point has been reached (Padilla & Benítez, 2014). To ensure theoretical saturation, the final sets of items were administered in three other classes of vocational high school ($n=68$).

Data analysis. The analysis of the cognitive interview data was conducted jointly between two members of the research team (GJ; DT). For each interview question, the two researchers coded impressions to determine whether the participant understood or not the item (Table 6). The analysis of the cognitive interviews data was twofold: (1) the analysis of each

misunderstood item and (2) the analysis of each item understood but justified by an unexpected explanation. Based on the problem encoded, the audio recording, and notes, the two experts worked together until they reached a consensus on the decision to be taken (i.e., discard, modify, or keep the item and reformulation). The researchers sought to determine whether an alternative formulation could improve comprehension and what the best possible wording would be (Supplementary file 5). In the case of non-consensus, a third member (PF) contributed to the consensus to identify the best rewording to test in the next step.

Results of Study 2

Each interview ($n=32$) was successfully completed without interruption. At the end of each stage, the researchers met to adjust the questionnaire according to the interviews (Supplementary file 5). The cognitive interview data allowed precise questionnaires according to: (1) rephrasing of items not understood or misunderstood based on the participants' answers (Table 6); and (2) modifying items that did not assess what they were supposed to (Table 6). At the end of this process, no item presented any misunderstanding and thus, the content validity of the three questionnaires (i.e., affective, cognitive, and social including the 8 aquatic items) was ensured. Significant modifications concerning the form of the questionnaire were made (i.e., item headers, highlighted).

At the end of the cognitive interview process, we noted any comments about understanding in the three vocational test classes, and no student refused to answer. The time required to complete the questionnaire was between 10 and 20 minutes for each document. At the end of this phase, we obtained an initial pool of 88 items for testing affective, cognitive, and social dimensions of the ELIP. These items were completed with three motor tests (Figure 1). Each test is presented in Appendix (translated into English for the reader's understanding).

[Table 6 near here].

Discussion

The aim of this study was to develop a measure to assess PL for emerging adults. To address these epistemological and methodological challenges, two consecutive studies were conducted: a Delphi method and a cognitive interview process. Through this process, we developed a comprehensive PL assessment tool useful with four dimensions of PL: physical, social, affective, and cognitive.

An original tool designed according to four dimensions. To our knowledge, ELIP is the first tool to simultaneously evaluate these four dimensions for emerging adults while, at the same time, considering the philosophical pillars of PL. Despite the social dimension being already included in robust definitions of PL (Keegan et al., 2019; Martins et al., 2020, Physical Literacy for Life, 2021), ELIP is one of very few PL assessment tools to consider this dimension (Barnett et al., 2020; Lodewyk, 2019; Mota et al., 2021). Some other tests refer to the social dimension (PFL, PPLI) but do not orient this dimension in the specific field of PA. The PL-C Quest and PPLA-Q were designed to map to the Australian Framework (Keegan et al., 2019) and do have a domain which focuses on the social, but are aimed at children and adolescents, respectively. This focus of the social dimension deviates from Whitehead's well-known definition and is probably a consequence of new approaches to PL, included the Australian and European approaches (Keegan et al., 2019; Physical Literacy for Life, 2021) and the specificity of the target population. For emerging adults, the social dimension seems to be a key dimension of the PA commitment process (Lu & Steele, 2014; Van Der Horst et al., 2007). Historically, inclusion of the social domain was first published by Dudley (2015) and early references to social domains were already present in the additional attributes of PL (including the ability to work independently and with others in both

cooperative and competitive situations, Whitehead, 2018). Our results converge with these approaches including the social dimension as an integral component of PL, and therefore invite its consideration in future studies and evaluation tools, especially for the emerging adult population for whom social relationships appear to be essential for quality of life (Edwards et al., 2002).

In view of the affective dimension, ELIP is in line with the previous proposals of ‘psychological dimension’ (e.g., CAPL, IPLA charting). The set of PL tests mentioned above assess this dimension. However, in the ELIP, the affective dimension is distinguished by considering the affective dimension in three different ways: affect towards PA (i.e., pleasure, well-being); affect in PA context (i.e., confidence, self-esteem); management of affect in PA context (i.e., management of emotions). ELIP is one of the recent tools [see also the PL-C Quest (Barnett et al., 2021)] which focus on the management of emotions in PA context.

The ELIP cognitive dimension focuses on individual knowledge and cognitive resources and does not deviate too widely from the tools already designed. The originality of the results lies in coupling between Likert scales assessment and a factual knowledge test. Here, the cognitive dimension goes beyond the fact of ‘knowing about PA’ but is also concerned with its application and personal awareness through a self-assessment. The importance of this dimension is already advocated (Cairney et al., 2019b; Cale & Harris, 2018) and is essential to consider for emerging adolescents insofar as it is poorly or not at all supported at school, at least in France (Gandrieau et al., 2021).

The results of the Delphi method are also original concerning different points of the physical dimension. First, ELIP includes on potential innovation in that the fitness dimension included in some measures (e.g., strength, endurance, etc.) is not explicitly present. Indeed, the fitness dimension was rejected by the experts, who emphasized that it is particularly subject to normative comparison and does not represent a fundamental asset for sustainable

physical activity education (Rowland, 1995). This perspective clearly differentiates the ELIP from the other tools which emphasize the importance of the fitness dimension (e.g., CAPL; PFL, Young et al., 2021). This choice is in line with some criticisms of pre-existing tools (Robinson & Randall, 2017), but opposes some defended views of PL (Tremblay & Longmuir, 2017). The expert panel focused their expectations on aspects of body control rather than physiological resources. It can be argued by its link with the generic term ‘literacy’ which can be understood as the ability to interact or engage effectively with the environment, in which we are situated with our own capabilities. Fitness, from a performance perspective, is not within the scope of most definitions of PL (Shearer et al., 2018). From this point of view, ELIP could be more consistent with the philosophy of the concept than other tests including fitness tests. In terms of the pedagogical aspect, this choice is significant. In fact, improving fitness can occur with attention to motor competence (Barnett et al., 2021) and ELIP invites PL professionals to lower the focus on fitness as a major goal but rather as one of many elements that would result from regular participation in PA. On this point, the skills needed to engage in resistance or fitness training could be questioned in further considerations. There are already tools to explore this important aspect for emerging adults (Lubans et al., 2014), but experts chose to stand at a more general level for the ELIP. It seems to imply that PL would be placed at a more global level.

The Delphi method introduces a different way of considering the motor tests of a PL assessment tool. Indeed, the *Star Challenge* test diverges from traditional Fundamental Motor Skills (FMS, Gallahue et al., 2012) by mobilizing broader movement competences in more complex and open environment (e.g., control body movements in an environment that must be decoded). Participants must perform motor skills (i.e., jumping, crawling, running, catching, etc.) and engage their movement abilities (i.e., stability, speed) jointly (Durdin-Myers et al., 2018) in environments, which challenge their interpretation. Although the anchoring of FMS

in the theory of constraint-led approach has been advocated (Barnett et al., 2016), the selected competences deviate from an isolated evaluation of FMS and emphasize the importance of “*action intelligence [...] embedded in perception, experience, memory, anticipation and decision making*” (Whitehead, 2001, p. 131). This specificity gives interesting anchorage with the existentialist perspective. To limit bias (i.e., anxiety, haste) and to best fit the concept of PL, the consideration of time as an evaluation criterion will have to be discussed in more detail according to the motor tasks performed during the test.

The novelty in the physical dimension is also found in the inclusion of cycling and aquatic competences already highlighted by the Australian approach (Barnett et al., 2020; Keegan et al., 2019), and the new perspectives to foundational movement skills (Hulteen et al., 2018). The specific reference to the aquatic and cycling world is relevant in a holistic French PL assessment tool. In fact, aquatic activities are culturally anchored in the practices of the French population, representing the 3rd most popular PA category (Croutte & Müller, 2018). Cycling also represents a huge opportunity for PA in France, with 18% of the population involved (Croutte & Müller, 2018). Moreover, both activities represent an opportunity for lifelong commitment in physical activity, as they can be practiced in France throughout life (Leger et al., 2019; Potdevin et al., 2013). Overall, these activities are also widely practiced in Europe and worldwide (Hulteen et al., 2017). These two specificities, which have largely not been considered in a PL test [see the PL-C Quest as an exception (Barnett et al., 2021)], connect PL with numerous beneficial PA opportunities for physical and mental health (Cox et al., 2010; Larouche et al., 2011). Nevertheless, experts seem to stand by a pragmatic position to make the assessment tool easily applicable. They made strong choices regarding water and cycling competences. Concerning aquatic competences, the experts chose to focus on perceived rather than objective ones. Therefore, attention must be paid to the gap between these two types of competences to avoid a poor-quality aquatic

education. Concerning cycling competences, while other wheeled transportations are available (e.g., scooters, skateboards, etc.) the experts chose to focus on bikes. Other active transportation competences should not be excluded from the development of physical domain.

Finally, the uniqueness of the physical dimension is also the selection of motor creativity, which has not yet been explored in PL assessment. Integrating the creativity dimension in a PL tool seems to correspond to the phenomenological perspective in which no precise model is expected but considers the uniqueness of the relationship between the individual and his/her environment that is missing from the existing PL's assessment tools, as well as ecological-dynamics framework perspectives (Rudd et al., 2020). Likewise, Davids et al. (2016) emphasized that creativity is an important component of adaptability. On this point, the expert consensus explicitly specified a new PL attribute converging towards the Whitehead's comprehensive understanding stating PL as "a disposition to use experience, understanding and abilities to interact effectively with the world" (Whitehead, 2004, p. 4).

These results of the Delphi method suggest ELIP is based on a higher level of abstraction on the Young et al. ladder (2021) than Whitehead's definition (2010). It exposes a conceptual understanding that emphasizes the importance of the social dimension and explicitly states new PL attributes considered necessary for the European French emerging adult audience. The results converge with the Physical Literacy for Life definition (2021) and Australian framework (Keegan et al., 2019), which emphasize the need to develop physical, emotional/affective, social, and cognitive skills and attributes. The Australian framework originally conceptualizes 32 PL elements, and we propose that ELIP could expand and operationalize the European definition to identify PL profiles in France and Europe. Despite an explicit stance on attributes, ELIP does not seem to lose sight of the epistemological foundations of the concept.

A tool designed according to the three philosophical pillars of PL. ELIP is a monitoring and evaluation tool consistent with Whitehead's philosophical perspective anchors. In fact, the existentialist perspective was respected by the wide nature of the PA definition and assessment environments (vs. sports-oriented vision and closed-motor pattern). Items were modified with the dual objectives of “*Physical activity should not be limited to sports (broad movement culture)*” and “*Adolescents should be able to easily project himself/herself into specific experiences that are unique to him/her*”. Moreover, the combination of questionnaires with complex objective tests allows capturing an interesting range of information for understanding the unique relationship between the individual and the environment. The complex nature of motor tests is particularly interesting in this respect of existentialism perspective. Now, there is new challenge in considering the singularity of the individual/environment relationship, particularly in the construction of scoring assessment. The complex system approaches (Preiser, 2019) could go beyond the limits announced by the ‘idealists’ (Edwards et al., 2018), opposed to PL assessment, by considering the diversity and uniqueness of PL patterns.

The anchoring in monist perspective is underlined with the holistic consideration of the whole being constituted by four dimensions (i.e., affective, cognitive, physical, and social). We argue that a tool design through four dimensions is both a methodological and epistemological requirement, but to truly grasp the concept as close as possible to its monistic underpinning there is a need to not slice off components so as to achieve a real integrated and monistic assessment. ELIP considers a broad multidimensional range of human components (Table 2) and the strength of the assessed concept resides in their interactions (Dudley, 2018). Then, to further embrace the challenge posed by this monistic perspective, a more complex scoring method needs to be explored to consider the interdependence and equal importance of each dimension.

Finally, the challenge of the phenomenological perspective must now be considered in the use of ELIP for the emerging adults to be properly addressed. Nevertheless, ELIP has the necessary structure for an accurate ipsative and formative assessment to inform the individuals' own PL journey. The tool will need to be used in this sense, which some may name as a 'charting' process (Whitehead, 2019), while further statistical analysis is needed to optimize its pragmatic form.

An empirical and practical tool. ELIP could provide a significant contribution both to research on PL and the practices that underpin it. It will allow measurement of PL in an original way during a critical transition of life (Arnett, 2000) by rooting it at a higher level of abstraction than Whitehead's well-known definition (Young et al., 2019). ELIP explicitly states about what and how to assess the PL attributes for emerging adults (e.g., motor creativity, cycling and aquatic competences, and management of emotions). This standing might enable practitioners' acceptance, while trying to preserve the philosophical perspectives which carry strong pedagogical values.

On one hand, ELIP might help teachers' guide and support students in their PL journeys by analyzing student profiles to identify the capabilities and competencies that need to be developed to access more PA opportunities throughout life. Missing or dodging a test is not a disqualifying fault, it will only be notified that the participant is not able, at this time, to meet the requirements of this test which therefore limits his/her opportunity for being active (e.g., cycling competences). This should be considered for the pedagogical or scientific analysis. Then, ELIP could be a key tool in PE curriculum, particularly in the last year of high school, which represents the last year of compulsory PE for the European French emerging adults. On the other hand, it will be useful to study the importance of PL and how it changes during the transition from adolescence to adulthood. The study of this life transition

is essential to complete the knowledge of PL value throughout the life course and to understand the PL journey (Longmuir & Tremblay, 2016).

Future research. A Delphi method with good engagement from PL experts, and the cognitive interviews with students from the targeted age-groups were useful to overcome the difficulties of creating a new educational assessment tool, and specifically to verify face and content validity, as well as assess feasibility. The results of this iterative method must now be tested by a proof of concept to confront the theoretical product to the real complexity of the concept. It would be interesting to organize a focus group based on the first quantitative results to clarify and confirm the tool being designed. The major challenge will be to finalize the tool with scoring method that best represents the PL concept. Likewise, the design of the tool should be finalized by checking the four methodological pillars: the feasibility, sensitivity, reliability, and validity.

Conclusions. Initially, we asked how to assess PL for emerging adults (17-21 years old). This study allowed the design of a new PL assessment tool conceptually validated by a significant number of international experts and an interview cognitive process. Without questioning the strength of validation of other widely used tools, ELIP complements the literature by providing an original and unique perspective to the field of PL assessment. ELIP is the first tool to investigate PL among emerging adults by combining subjective and motor tests. Its four interdependent dimensions – physical, social, cognitive, and affective – emphasize an explicit wide range of holistic resources (Table 2) that broadens considerably the affordances an individual may be able to grasp (O’Sullivan et al., 2020). The design methodology retained the philosophical anchoring of the tool, but the validation process and progress on feasibility still need to be further measured for a real implementation plan for educators and teachers. Our results contribute to the improvement of the PL measurement

tools literature and, subsequently, participate in the promotion and monitoring of PL among emerging adults.

References

- Alipour Anbarani, M., Ghaffari, M., Montazeri, ali, Kavousi, A., & Ramezankhani, A. (2022). Physical Literacy of 16-18-Years Adolescents: A Qualitative Study. *International Journal of Pediatrics*, 10(1). <https://doi.org/10.22038/ijp.2021.60209.4671>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Arnett, J. J. (2000). Emerging adulthood. A theory of development from the late teens through the twenties. *The American Psychologist*, 55(5), 469–480.
- Barnett, L. M., Dudley, D. A., Telford, R. D., Lubans, D. R., Bryant, A. S., Roberts, W. M., Morgan, P. J., Schranz, N. K., Weissensteiner, J. R., Vella, S. A., Salmon, J., Ziviani, J., Okely, A. D., Wainwright, N., Evans, J. R., & Keegan, R. J. (2019). Guidelines for the Selection of Physical Literacy Measures in Physical Education in Australia. *Journal of Teaching in Physical Education*, 38(2), 119–125. <https://doi.org/10.1123/jtpe.2018-0219>
- Barnett, L. M., Mazzoli, E., Hawkins, M., Lander, N., Lubans, D. R., Caldwell, S., Comis, P., Keegan, R. J., Cairney, J., Dudley, D., Stewart, R. L., Long, G., Schranz, N., Brown, T. D., & Salmon, J. (2020). Development of a self-report scale to assess children’s perceived physical literacy. *Physical Education and Sport Pedagogy*, 1–26. <https://doi.org/10.1080/17408989.2020.1849596>
- Barnett, L. M., Stodden, D., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., Iivonen, S., Miller, A. D., Laukkanen, A., Dudley, D., Lander, N. J., Brown, H., & Morgan, P. J. (2016). Fundamental Movement Skills: An Important Focus. *Journal of Teaching in Physical Education*, 35(3), 219–225. <https://doi.org/10.1123/jtpe.2014-0209>
- Barnett, L. M., Webster, E. K., Hulteen, R. M., De Meester, A., Valentini, N. C., Lenoir, M., Pesce, C., Getchell, N., Lopes, V. P., Robinson, L. E., Brian, A., & Rodrigues, L. P. (2021). Through the Looking Glass: A Systematic Review of Longitudinal Evidence, Providing New Insight

- for Motor Competence and Health. *Sports Medicine*. <https://doi.org/10.1007/s40279-021-01516-8>
- Bauman, A. E., Reis, R. S., Sallis, J. F., Wells, J. C., Loos, R. J., & Martin, B. W. (2012). Correlates of physical activity: Why are some people physically active and others not? *The Lancet*, *380*(9838), 258–271. [https://doi.org/10.1016/S0140-6736\(12\)60735-1](https://doi.org/10.1016/S0140-6736(12)60735-1)
- Beatty, P. C., & Willis, G. B. (2007). Research Synthesis: The Practice of Cognitive Interviewing. *Public Opinion Quarterly*, *71*(2), 287–311. <https://doi.org/10.1093/poq/nfm006>
- Berndt, T. J., & Savin-Williams, R. C. (1993). Peer relations and friendships. In *Handbook of clinical research and practice with adolescents*. (pp. 203–219). John Wiley & Sons.
- Bopp, T., & Vadeboncoeur, J. D. (2019). *Assessing Affective Physical Literacy of Adolescents: The Development of a Motivation and Confidence in Sport and Physical Activity Instrument*. *6*(2), 56–69. <https://doi.org/10.15640/jpesm.v6n2a7>
- Brasseur, S., Grégoire, J., Bourdu, R., & Mikolajczak, M. (2013). The Profile of Emotional Competence (PEC): Development and Validation of a Self-Reported Measure that Fits Dimensions of Emotional Competence Theory. *PLoS ONE*, *8*(5), e62635. <https://doi.org/10.1371/journal.pone.0062635>
- Cairney, J., Clark, H. J., James, M. E., Mitchell, D., Dudley, D. A., & Kriellaars, D. (2018). The Preschool Physical Literacy Assessment Tool: Testing a New Physical Literacy Tool for the Early Years. *Frontiers in Pediatrics*, *6*, 138. <https://doi.org/10.3389/fped.2018.00138>
- Cairney, J. (2019a, September). The Science Turn in Physical Literacy: is the field shifting and why?. [Presentation]. International Physical Literacy Conference, UMEA, Sweden.
- Cairney, J., Dudley, D., Kwan, M., Bulten, R., & Kriellaars, D. (2019b). Physical Literacy, Physical Activity and Health: Toward an Evidence-Informed Conceptual Model. *Sports Medicine*, *49*(3), 371–383. <https://doi.org/10.1007/s40279-019-01063-3>
- Caldwell, H. A., Di Cristofaro, N. A., Cairney, J., Bray, S. R., & Timmons, B. W. (2021). Measurement properties of the Physical Literacy Assessment for Youth (PLAY) Tools. *Applied Physiology, Nutrition, and Metabolism*, *46*(6), 571–578. <https://doi.org/10.1139/apnm-2020-0648>

- Cale, L., & Harris, J. (2018). The Role of Knowledge and Understanding in Fostering Physical Literacy. *Journal of Teaching in Physical Education*, 37(3), 280–287.
<https://doi.org/10.1123/jtpe.2018-0134>
- Chen, A. (2020a). A clash of fundamental assumptions: Can/should we measure physical literacy? *Journal of Sport and Health Science*, 9(2), 149–151.
<https://doi.org/10.1016/j.jshs.2019.11.002>
- Chen, S.-T., Tang, Y., Chen, P.-J., & Liu, Y. (2020b). The Development of Chinese Assessment and Evaluation of Physical Literacy (CAEPL): A Study Using Delphi Method. *International Journal of Environmental Research and Public Health*, 17(8), 2720.
<https://doi.org/10.3390/ijerph17082720>
- Cox, K. L., Burke, V., Beilin, L. J., & Puddey, I. B. (2010). A comparison of the effects of swimming and walking on body weight, fat distribution, lipids, glucose, and insulin in older women—The Sedentary Women Exercise Adherence Trial 2. *Metabolism*, 59(11), 1562–1573.
<https://doi.org/10.1016/j.metabol.2010.02.001>
- Croutte, P., & Müller, J. (2018). *Baromètre national des pratiques sportives 2018* [Baromètre réalisé par le CREDOC sous la direction de Hoibian S. pour l'INJEP et le ministère des sports]. INJEP Notes & rapports/Rapport d'étude.
- Davids, K., Araújo, D., & Brymer, E. (2016). Designing Affordances for Health-Enhancing Physical Activity and Exercise in Sedentary Individuals. *Sports Medicine*, 46(7), 933–938.
<https://doi.org/10.1007/s40279-016-0511-3>
- Dietrich, H., & Ehrlenspiel, F. (2010). Cognitive Interviewing: A Qualitative Tool for Improving Questionnaires in Sport Science. *Measurement in Physical Education and Exercise Science*, 14, 51–60. <https://doi.org/10.1080/10913670903455025>
- Dudley, D. (2018). Physical Literacy: When the Sum of the Parts Is Greater than the Whole. *Journal of Physical Education, Recreation & Dance*, 89(3), 7–8.
<https://doi.org/10.1080/07303084.2018.1418998>
- Dudley, D. A. (2015). A Conceptual Model of Observed Physical Literacy. *The Physical Educator*.
<https://doi.org/10.18666/TPE-2015-V72-I5-6020>

- Durden-Myers, E. J., Green, N. R., & Whitehead, M. E. (2018). Implications for Promoting Physical Literacy. *Journal of Teaching in Physical Education*, 37(3), 262–271.
- Edwards, L. C., Bryant, A. S., Keegan, R. J., Morgan, K., Cooper, S.-M., & Jones, A. M. (2018). ‘Measuring’ Physical Literacy and Related Constructs: A Systematic Review of Empirical Findings. *Sports Medicine*, 48(3), 659–682. <https://doi.org/10.1007/s40279-017-0817-9>
- Edwards, L. C., Bryant, A. S., Keegan, R. J., Morgan, K., & Jones, A. M. (2017). Definitions, Foundations and Associations of Physical Literacy: A Systematic Review. *Sports Medicine*, 47(1), 113–126. <https://doi.org/10.1007/s40279-016-0560-7>
- Edwards, T. C., Huebner, C. E., Connell, F. A., & Patrick, D. L. (2002). Adolescent quality of life, Part I: Conceptual and measurement model. *Journal of Adolescence*, 25(3), 275–286. <https://doi.org/10.1006/jado.2002.0470>
- Fox, K. R., & Corbin, C. B. (1989). The Physical Self-Perception Profile: Development and Preliminary Validation. *Journal of Sport and Exercise Psychology*, 11(4), 408–430. <https://doi.org/10.1123/jsep.11.4.408>
- Gallahue, D., Ozmun, J., & Goodway, J. (2012). *Understanding Motor Development: Infants, Children, Adolescents, Adults. Seventh edn. (7th ed.)*.
- Gandrieau, J., Schnitzler, C., Derigny, T., Ll ena, C., Mouton, A., & Potdevin, F. (2021). How Knowledge About Physical Activity Is Impacted By School Institution, Grade Level, and Gender Throughout High School Years in France? *Journal of Teaching in Physical Education*, 1–9. <https://doi.org/10.1123/jtpe.2021-0092>
- Green, N. R., Roberts, W. M., Sheehan, D., & Keegan, R. J. (2018). Charting Physical Literacy Journeys Within Physical Education Settings. *Journal of Teaching in Physical Education*, 37(3), 272–279. <https://doi.org/10.1123/jtpe.2018-0129>
- Gresham, F. M., & Elliott, N. (2008). *Social Skills Improvement System Rating Scales*. (NCS Pearson).
- Gunnell, K. E., Schellenberg, B. J. I., Wilson, P. M., Crocker, P. R. E., Mack, D. E., & Zumbo, B. D. (2014). A Review of Validity Evidence Presented in the Journal of Sport and Exercise Psychology (2002–2012): Misconceptions and Recommendations for Validation Research. In B. D. Zumbo & E. K. H. Chan (Eds.), *Validity and Validation in Social, Behavioral, and*

- Health Sciences* (Vol. 54, pp. 137–156). Springer International Publishing.
https://doi.org/10.1007/978-3-319-07794-9_8
- Hanson, C. L., Oliver, E. J., Dodd-Reynolds, C. J., Pearsons, A., & Kelly, P. (2020). A modified Delphi study to gain consensus for a taxonomy to report and classify physical activity referral schemes (PARS). *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 158. <https://doi.org/10.1186/s12966-020-01050-2>
- Harter, S. (2017). *Self-Perception Profile for Adolescents* [Data set]. American Psychological Association. <https://doi.org/10.1037/t05703-000>
- Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*(3), 238–247. <https://doi.org/10.1037/1040-3590.7.3.238>
- Heino, M. T. J., Knittle, K., Noone, C., Hasselman, F., & Hankonen, N. (2021). Studying Behaviour Change Mechanisms under Complexity. *Behavioral Sciences*, *11*(5), 77.
<https://doi.org/10.3390/bs11050077>
- Hochberg, Z., & Konner, M. (2020). Emerging Adulthood, a Pre-adult Life-History Stage. *Frontiers in Endocrinology*, *10*, 918. <https://doi.org/10.3389/fendo.2019.00918>
- Hulteen, R. M., Morgan, P. J., Barnett, L. M., Stodden, D. F., & Lubans, D. R. (2018). Development of Foundational Movement Skills: A Conceptual Model for Physical Activity Across the Lifespan. *Sports Medicine*, *48*(7), 1533–1540. <https://doi.org/10.1007/s40279-018-0892-6>
- Hulteen, R. M., Smith, J. J., Morgan, P. J., Barnett, L. M., Hallal, P. C., Colyvas, K., & Lubans, D. R. (2017). Global participation in sport and leisure-time physical activities: A systematic review and meta-analysis. *Preventive Medicine*, *95*, 14–25.
<https://doi.org/10.1016/j.ypmed.2016.11.027>
- IPLA. (2017). *IPLA definition*. Retrieved from. <https://www.physical-literacy.org.uk/>
- Jean de Dieu, H., & Zhou, K. (2021). Physical Literacy Assessment Tools: A Systematic Literature Review for Why, What, Who, and How. *International Journal of Environmental Research and Public Health*, *18*(15), 7954. <https://doi.org/10.3390/ijerph18157954>

- Ke, D., Suzuki, K., Kishi, H., Kurokawa, Y., & Shen, S. (2022). Definition and assessment of physical literacy in children and adolescents: A literature review. *The Journal of Physical Fitness and Sports Medicine*, *11*(3), 149–159. <https://doi.org/10.7600/jpfsm.11.149>
- Keegan, R. J., Barnett, L. M., Dudley, D. A., Telford, R. D., Lubans, D. R., Bryant, A. S., Roberts, W. M., Morgan, P. J., Schranz, N. K., Weissensteiner, J. R., Vella, S. A., Salmon, J., Ziviani, J., Okely, A. D., Wainwright, N., & Evans, J. R. (2019). Defining Physical Literacy for Application in Australia: A Modified Delphi Method. *Journal of Teaching in Physical Education*, *38*(2), 105–118. <https://doi.org/10.1123/jtpe.2018-0264>
- Kendzierski, D., & DeCarlo, K. J. (1991). Physical Activity Enjoyment Scale: Two Validation Studies. *Journal of Sport and Exercise Psychology*, *13*(1), 50–64. <https://doi.org/10.1123/jsep.13.1.50>
- Kwan, M. Y. W., Graham, J. D., Bedard, C., Bremer, E., Healey, C., & Cairney, J. (2019). Examining the Effectiveness of a Pilot Physical Literacy–Based Intervention Targeting First-Year University Students: The PLUS Program. *SAGE Open*, *9*(2), 215824401985024. <https://doi.org/10.1177/2158244019850248>
- Larouche, R., Lloyd, M., Knight, E., & Tremblay, M. S. (2011). Relationship Between Active School Transport and Body Mass Index in Grades—4-to-6 Children. *Pediatric Exercise Science*, *23*(3), 322–330.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Leger, S. J., Dean, J. L., Edge, S., & Casello, J. M. (2019). “If I had a regular bicycle, I wouldn’t be out riding anymore”: Perspectives on the potential of e-bikes to support active living and independent mobility among older adults in Waterloo, Canada. *Transportation Research Part A: Policy and Practice*, *123*, 240–254. <https://doi.org/10.1016/j.tra.2018.10.009>
- Liu, C.-Y., Lin, L. L.-C., Sheu, J.-J., & Sum, R. K.-W. (2022). Psychometric Validation of Senior Perceived Physical Literacy Instrument. *International Journal of Environmental Research and Public Health*, *19*(11), 6726. <https://doi.org/10.3390/ijerph19116726>

- Lodewyk, K. R. (2019). Early Validation Evidence of the Canadian Practitioner-Based Assessment of Physical Literacy in Secondary Physical Education. *Physical Educator*, 76(3), 634–660.
- Longmuir, P. E., Boyer, C., Lloyd, M., Yang, Y., Boiarskaia, E., Zhu, W., & Tremblay, M. S. (2015). The Canadian Assessment of Physical Literacy: Methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Health*, 15(1), 767. <https://doi.org/10.1186/s12889-015-2106-6>
- Longmuir, P. E., & Tremblay, M. S. (2016). Top 10 Research Questions Related to Physical Literacy. *Research Quarterly for Exercise and Sport*, 87(1), 28–35. <https://doi.org/10.1080/02701367.2016.1124671>
- Lu, C., & Steele, K. (2014). Characteristics and conceptual framework of the Easy-Play Model. *European Physical Education Review*, 20(2), 229–243. <https://doi.org/10.1177/1356336X14524861>
- Lubans, D. R., Smith, J. J., Harries, S. K., Barnett, L. M., & Faigenbaum, A. D. (2014). Development, test-retest reliability, and construct validity of the resistance training skills battery. *Journal of Strength and Conditioning Research*, 28(5), 1373–1380. <https://doi.org/10.1519/JSC.0b013e31829b5527>
- Markland, D., & Tobin, V. (2004). A Modification to the Behavioural Regulation in Exercise Questionnaire to Include an Assessment of Amotivation. *Journal of Sport and Exercise Psychology*, 26(2), 191–196. <https://doi.org/10.1123/jsep.26.2.191>
- Martins, J., Onofre, M., Mota, J., Murphy, C., Repond, R.-M., Vost, H., Cremosini, B., Svrdlim, A., Markovic, M., & Dudley, D. (2020). International approaches to the definition, philosophical tenets, and core elements of physical literacy: A scoping review. *PROSPECTS*. <https://doi.org/10.1007/s11125-020-09466-1>
- Méndez–Martínez, E., & Fernández–Río, J. (2019). Diseño Y Validación De Un Instrumento De Medición De La Creatividad Motriz En Adolescentes. *Revista Internacional de Medicina y Ciencias de La Actividad Física y Del Deporte*, 19(75), 535. <https://doi.org/10.15366/rimcafd2019.75.010>
- Merleau-Ponty, M. (1968). *Phenomenology of perception* (C. Smith, Trans.). Routledge.

- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research, 19*(4), 539–549.
<https://doi.org/10.1007/s11136-010-9606-8>
- Moran, K., Stallman, R. K., Kjendlie, P.-L., Dahl, D., Blitvich, J. D., Petrass, L. A., McElroy, G. K., Goya, T., Teramoto, K., Matsui, A., & Shimongata, S. (2012). Can You swim? An Exploration of Measuring Real and Perceived Water Competency. *International Journal of Aquatic Research and Education, 6*(2). <https://doi.org/10.25035/ijare.06.02.04>
- Mota, J., Martins, J., & Onofre, M. (2021). Portuguese Physical Literacy Assessment Questionnaire (PPLA-Q) for adolescents (15–18 years) from grades 10–12: Development, content validation and pilot testing. *BMC Public Health, 21*(1), 2183. <https://doi.org/10.1186/s12889-021-12230-5>
- Muller, J. (2018). *Les jeunes aiment le sport...de préférence sans contraintes* (No. 297; p. 3). UCPA/CREDOC.
- O’Sullivan, M., Davids, K., Woods, C. T., Rothwell, M., & Rudd, J. (2020). Conceptualizing Physical Literacy within an Ecological Dynamics Framework. *Quest, 1–15*.
<https://doi.org/10.1080/00336297.2020.1799828>
- Padilla, J.-L., & Benítez, I. (2014). Validity evidence based on response processes. *Psicothema, 26.1*, 136–144. <https://doi.org/10.7334/psicothema2013.259>
- Physical Literacy For Life. (2021). <https://physical-literacy.isca.org>
- Pot, N., Whitehead, M. E., & Durden-Myers, E. J. (2018). Physical Literacy From Philosophy to Practice. *Journal of Teaching in Physical Education, 37*(3), 246–251.
<https://doi.org/10.1123/jtpe.2018-0133>
- Potdevin, F., Normani, C., & Pelayo, P. (2013). Examining Self-Training Procedures in Leisure Swimming. *Journal of Sports Science & Medicine, 12*, 716–723.
- Powell, C. (2003). The Delphi technique: Myths and realities. *Journal of Advanced Nursing, 41*(4), 376–382. <https://doi.org/10.1046/j.1365-2648.2003.02537.x>

- Preiser, R. (2019). Identifying general trends and patterns in complex systems research: An overview of theoretical and practical implications. *Systems Research and Behavioral Science*, 36(5), 706–714. <https://doi.org/10.1002/sres.2619>
- Robinson, D. B., & Randall, L. (2017). Marking Physical Literacy or Missing the Mark on Physical Literacy? A Conceptual Critique of Canada’s Physical Literacy Assessment Instruments. *Measurement in Physical Education and Exercise Science*, 21(1), 40–55. <https://doi.org/10.1080/1091367X.2016.1249793>
- Rowland, T. W. (Ed.). (1995). “The Horse Is Dead; Let’s Dismount.” *Pediatric Exercise Science*, 7(2), 117–120. <https://doi.org/10.1123/pes.7.2.117>
- Rudd, J. R., Pesce, C., Strafford, B. W., & Davids, K. (2020). Physical Literacy - A Journey of Individual Enrichment: An Ecological Dynamics Rationale for Enhancing Performance and Physical Activity in All. *Frontiers in Psychology*, 11, 1904. <https://doi.org/10.3389/fpsyg.2020.01904>
- Sallis, J. F., Pinski, R. B., Grossman, R. M., Patterson, T. L., & Nader, P. R. (1988). The development of self-efficacy scales for healthrelated diet and exercise behaviors. *Health Education Research*, 3(3), 283–292. <https://doi.org/10.1093/her/3.3.283>
- Shearer, C., Goss, H. R., Edwards, L. C., Keegan, R. J., Knowles, Z. R., Boddy, L. M., Durden-Myers, E. J., & Foweather, L. (2018). How Is Physical Literacy Defined? A Contemporary Update. *Journal of Teaching in Physical Education*, 37(3), 237–245. <https://doi.org/10.1123/jtpe.2018-0136>
- Sum, R. K. W., Ha, A. S. C., Cheng, C. F., Chung, P. K., Yiu, K. T. C., Kuo, C. C., Yu, C. K., & Wang, F. J. (2016). Construction and Validation of a Perceived Physical Literacy Instrument for Physical Education Teachers. *PLOS ONE*, 11(5), e0155610. <https://doi.org/10.1371/journal.pone.0155610>
- Tremblay, M. S., & Longmuir, P. E. (2017a). Conceptual Critique of Canada’s Physical Literacy Assessment Instruments Also Misses the Mark. *Measurement in Physical Education & Exercise Science*, 21(3), 174–176.

- Tremblay, M. S., & Longmuir, P. E. (2017b). Conceptual Critique of Canada's Physical Literacy Assessment Instruments Also Misses the Mark. *Measurement in Physical Education and Exercise Science*, 21(3), 174–176. <https://doi.org/10.1080/1091367X.2017.1333002>
- Van Der Horst, K., Paw, M. J. C. A., Twisk, J. W. R., & Van Mechelen, W. (2007). A brief review on correlates of physical activity and sedentariness in youth. *Medicine and Science in Sports and Exercise*, 39(8), 1241–1250. <https://doi.org/10.1249/mss.0b013e318059bf35>
- Wang, C. K. J., Liu, W. C., Biddle, S. J. H., & Spray, C. M. (2005). Cross-cultural validation of the Conceptions of the Nature of Athletic Ability Questionnaire Version 2. *Personality and Individual Differences*, 38(6), 1245–1256. <https://doi.org/10.1016/j.paid.2004.08.007>
- Whitehead, M. (2001). The Concept of Physical Literacy. *European Journal of Physical Education*, 6(2), 127–138. <https://doi.org/10.1080/1740898010060205>
- Whitehead, M. (2007). Physical Literacy: Philosophical Considerations in Relation to Developing a Sense of Self, Universality and Propositional Knowledge. *Sport, Ethics and Philosophy*, 1(3), 281–298. <https://doi.org/10.1080/17511320701676916>
- Whitehead, M. (2010). *Physical Literacy: Throughout the Lifecourse*. Routledge. <https://doi.org/10.4324/9780203881903>
- Whitehead, M. (2019). *Physical Literacy across the World*. Routledge. <https://doi.org/10.4324/9780203702697>
- Whitehead, M. E. (2004). physical literacy - a debate. (abstract). in *in Klisouras, v. (ed.), 2004 pre-olympic congress: proceedings: volume i: lectures-orals: sport science through the ages, 6-11 august 2004, Thessaloniki/hellas, Aristotle University campus, Thessaloniki, Greece, Aristotle University of Thessaloniki, Department of Physical Education & Sport Science, [2004], p.117-118.*
- Willis, G. B. (2005). *Cognitive Interviewing*. SAGE Publications Inc. <https://us.sagepub.com/en-us/nam/cognitive-interviewing/book225856>
- Young, L., O'Connor, J., & Alfrey, L. (2019). Physical literacy: A concept analysis. *Sport, Education and Society*, 25(8), 946–959. <https://doi.org/10.1080/13573322.2019.1677586>

Young, L., O'Connor, J., Alfrey, L., & Penney, D. (2021). Assessing physical literacy in health and physical education. *Curriculum Studies in Health and Physical Education*, 12(2), 156–179.
<https://doi.org/10.1080/25742981.2020.1810582>

Zhu, W. (2012). *Measurement Practice in Sport and Exercise Psychology: A Historical, Comparative, and Psychometric View*. <https://doi.org/10.5040/9781492596332.ch-002>