PHYSICAL LITERACY ASSESSMENT AMONG PRIMARY SCHOOL CHILDREN AGED 7-11 YEARS

Cara Shearer

A thesis submitted in partial fulfilment of the requirements of

Liverpool John Moores University for the degree of Doctor of

Philosophy

October 2019

Abstract

Physical literacy (PL) is the motivation, confidence, physical competence, knowledge and understanding to value and participate in physical activity for life. Assessment of physical literacy has recently emerged as an aspect of the concept that requires further investigation. Without the ability to measure PL, practitioners, teachers and policy makers cannot ascertain if children are displaying positive PL behaviours or if additional support is required. It is currently unclear if there are assessment tools available in research or practice that have reported philosophical alignment to Whitehead's concept of physical literacy and the development of one is pivotal to promote the operationalisation of the concept within the education sector.

Study One

Globally, PL continues to gain momentum, yet the definition and underlying concept of physical literacy remains contested in both research and practice. This lack of clarity has the potential to undermine the operationalisation of PL, as such study 1 considers the various definitions of PL that are currently adopted internationally. PL experts identified seven leading groups that have established PL initiatives. Although each group is unified in using the term PL, there are contrasting definitions and interpretations of the concept. Common themes were identified, including the: (a) influence of PL philosophy, (b) core elements of PL, (c) lifelong nature of PL, and (d) the need to scientifically pursue a robust operationalization of the concept. Study 1 concludes by recommending that programmes relating to PL should provide a definition, a clear philosophical approach, and transparency with how their actions align with this approach.

Study Two

The purpose of this study was to systematically review the literature for assessment tools that are appropriate to measure PL elements within children aged 7-11 years. A systematic review

was conducted in accordance with PRISMA-P guidelines. Search terms were defined during workshops with PL experts before the following electronic databases were searched (12 May 2017- 10 January 2019) to identify relevant peer-reviewed journal articles published in English: (i) MEDLINE (ii) ScienceDirect (iii) SPORTDiscus, iv) Education Research Complete (iv)Scopus, and (v) psycINFO. Methodological quality of both quantitative and qualitative assessment tools were appraised using the COSMIN risk of bias checklist. The feasibility, interpretability, cost and ease of use of each instrument was also appraised using a utility matrix and a bespoke PL checklist was developed to appraise alignment to the concept. The search strategy resulted in a total of 7530 initial results. Reference checking resulted in three additional eligible studies. After screening of titles and abstracts, 381 articles were retrieved for full text reading. After full text screening was completed a total of 130 studies were included in this review. In total, 65 studies were relevant to the physical domain, 58 to the affective domain and 7 to the cognitive domain. In relation to the 7-11 age range 21 tools were available for the physical domain, 33 for the affective and 6 for the cognitive. The findings within the review highlight that aspects of the physical, cognitive and affective domains are currently being assessed, using tools that have acceptable psychometric properties. While assessments are typically proven to be feasible within a school context, further empirical research is needed to consider the feasibility of the scoring and administration of assessment tools by teachers as opposed to research teams.

Study 3

Study 3 explored stakeholder perceptions of a PL assessment for children aged 7-11 years. Specifically, this study aimed to explore key stakeholders' views of current practice, effective implementation and future directions of PL assessment to inform the development of a PL assessment for primary school-aged children. Purposive samples of children aged 7-11 years (10 focus groups, n=57), primary school teachers (6 focus groups, n=23) and experts in

physical literacy (3 focus groups, academics n=13, practitioners n=8) were recruited to take part in a series of concurrent focus groups. A semi-structured focus group guide was developed focusing on acceptability, demand and implementation of PL assessment. Focus groups were audio recorded, transcribed verbatim and analysed using inductive and deductive thematic analysis, with key themes organised into pen profiles. It was found that stakeholders viewed the assessment of physical literacy as important but reported that it was not currently a priority in many schools, resulting in variable practice. In addition, child responses centred on a desire for enjoyment/fun within the assessment experience and teachers recommended that assessment should be time-efficient, simple and useful. Experts advocated the use of longitudinal assessment strategies. Moreover, all stakeholders proposed using technology and self-assessment/reflection to support PL assessment. Although all stakeholders recognised a demand for PL assessment at the individual and population level, existing assessments did not meet their perceived needs. Future assessments should consider the balance between the purpose of the assessment and the feasibility of the assessment process.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Publications and Conference Outputs from the PhD

Scientific publications

- **Shearer**, C., Goss, H. R., Edwards, L. C., Keegan, R. J., Knowles, Z. R., Boddy, L. M., ... & Foweather, L. (2018). How is PL defined? A contemporary update. Journal of Teaching in PE, 37(3), 237-245.

- **Shearer**, C., Goss, H. R., Edwards, L. C., Keegan, R. J., Knowles, Z. R., Boddy, L. M., ... & Foweather, L. (under review). A systematic review of assessments related to the physical domain of physical literacy among children aged 3-11 years

Conference presentations (oral)

- **Shearer, C.,** Goss, H. R., Durden-Myers, E., Knowles, Z. R., Boddy, L. M., & Foweather, L. Physical Literacy. Physical Activity Exchange Seminar Series, Liverpool, January 2018

- **Shearer, C.,** Goss, H. R., Durden-Myers, E., Knowles, Z. R., Boddy, L. M., & Foweather, L. A Systematic Review of Assessments related to Physical Literacy. International Physical Literacy Conference, Liverpool and Cardiff June 2017 and June 2018

- **Shearer, C.,** Goss, H. R., Durden-Myers, E., Knowles, Z. R., Boddy, L. M., & Foweather, L. A Systematic Review of Assessments related to Physical Literacy. International Society for Physical Activity and Health, London, October 2018

- **Shearer, C**., Goss, H. R., Durden-Myers, E., Knowles, Z. R., Boddy, L. M., & Foweather, L. A Systematic Review of Assessments related to Physical Literacy. Motor Competence Research Day, Cork, January 2019

- **Shearer, C.,** Goss, H. R., Durden-Myers, E., Knowles, Z. R., Boddy, L. M., & Foweather, L. Engaging Stakeholder in Research. Science in Public Conference, Manchester, July 2019

- **Shearer, C.,** Goss, H. R., Durden-Myers, E., Knowles, Z. R., Boddy, L. M., & Foweather, L. Stakeholder Perspectives of Physical Literacy Assessment. Health-Enhancing Physical Activity, Odense, August 2019

- **Shearer, C.,** Goss, H. R., Durden-Myers, E., Knowles, Z. R., Boddy, L. M., & Foweather, L. A Systematic Review of Assessments related to Physical Literacy; Stakeholder Perspectives of Physical Literacy Assessment. International Physical Literacy Conference (Sport for Life), Umea, September 2019

- **Shearer, C.,** Goss, H. R., Durden-Myers, E., Knowles, Z. R., Boddy, L. M., & Foweather, L. A Systematic Review of Assessments related to Physical Literacy; Stakeholder Perspectives of Physical Literacy Assessment. Scottish Physical Activity Research Conference, Edinburgh, October 2019

Acknowledgements

Firstly, I'd like to thank the fantastic team within the Physical Activity Exchange. I have had such a positive and enriching experience within the department and it has been a privilege to learn from such an inspiring group of students and staff. I'd like to thank my supervisory team Dr Lawrence Foweather, Dr Lynne Boddy, Professor Zoe Knowles and Elizabeth Durden-Myers for being so generous with your time and providing constant support both on an academic and personal level. Special thanks to my Director of Studies, Lawrence for being an excellent mentor, always leading by example and going above and beyond what is expected. Thanks also to co-researcher Hannah Goss for being a great colleague and friend.

To my wonderful family, thank you for always rallying around when I need it most and reminding me what is important in life. Mum and Dad, I cannot thank you enough for the constant support and guidance- I hope one day I can support Noah to the lengths that you have supported me. Continuing my education to PhD level is a privilege. I would not have been able to do so without the family members who have gone before me, who worked tirelessly to give me to opportunities that I have today- thank you.

The greatest thanks must go to my wife and partner Karen. This would not have been possible without your unwavering patience, support and love. Most importantly, thank you for believing in me on the days where I didn't believe in myself.

The work within this thesis is dedicated to my son Noah Shearer who was born in the final stages of this project (09/06/2019). Both you and your mum continue to inspire me every day and I hope that in the future you have the pleasure of following your passion.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
IPLA	International Physical Literacy Association
PA	Physical Activity
PE	Physical Education
PL	Physical Literacy
UK	United Kingdom
USA	United States of America
CS4L	Canadian Sport for Life
PHE Canada	Physical and Health Education Canada
PLFL	Physical Literacy for Life
SHAPE	Society of Health and Physical Educators
COSMIN	<mark>COnsensus-based</mark> Standards for the selection of health Measurement Instruments
PROM	Patient Reported Outcome Measure
NR	Not reported
ALPHA	ALPHA Fitness Battery
AST	Athletic Skills Track
BOTMP-SF	Bruininks–Oseretsky Test of Motor Proficiency
CAMSA	Canadian Agility and Movement Skills Assessment
CAPL-2	Canadian Assessment of Physical Literacy (2)
EUROFIT	The Eurofit Physical Fitness Test Battery
FGCOMP	FG-COMPASS
FG	FITNESSGRAM
GSPA	Golf Swing and Putt skill Assessment

HS	High/Scope beat competence analysis test
KTK	Körperkoordinationstest für Kinder
MOBAK-3	Motorische Basiskompetenzen in der 3
MABC2	Movement assessment battery for children-2
MUGI	Motorisk Utveckling som Grund för Inlärning
OP	Obstacle Polygon
PARAGON	PA Research and Assessment tool for Garden Observation
PF	PLAYfun
SMT	Slalom Movement Test
SEBT	Star Excursion Balance Test
TGMD-3	Test of Gross Motor Development-3
20MSR	The Leger 20m Shuttle Run test
YBT	Y Balance Test
FHC-Q	Food, Health, and Choices Questionnaire
ATOP	Attitudes Towards Outdoor Play Scales
AGSYS	Achievement Goal Scale for Youth Sports
CPAS	Commitment to PA Scale
LEAP	Lunchtime Enjoyment Activity and Play
PASE	PA Self-Efficacy Questionnaire
ATCPE	Attitudes Towards Curriculum PE Scale
RCS	Response to Challenge Scale
PACES	PA Enjoyment Scale
CY-PSPP	The Children and Youth Self-Perception Profile
CMPI	Children's Perception of Motor Competence Scale

PAHFE	PA and Healthy Food Efficacy
MOSS	Motivational Orientation in Sport Questionnaire
CAPA	Children's Attraction to PA
HOP'N Evaluation	Healthy Opportunities for PA and Nutrition
PASES	PA Self- Efficacy Scales
SPPC	Self-Perception Profile for Children
CATPA	Children's Attitudes toward PA
DPAPI	Determinants of PA Participatory Inventory
PMCSQ	Perceived Motivational Climate in Sport Questionnaire
Pre-FPQ	Preschool PA and Food Questionnaire
BONES-PAS	Beat Osteoporosis Now-PA Survey
PH-KA	Pupils Health Knowledge Assessment
SCALES	Scales to measure knowledge and preference of diet and PA

Glossary of Terms

Physical activity	Defined as *"any bodily movement produced by skeletal muscles resulting in energy expenditure"* (Casperson et al., 1985, p.126)
Physical literacy	Defined as *"- the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for engagement in physical activities for life,""* (Whitehead, 2019, p.8)
Children	For the purpose of this thesis, the term children refers to those aged between 7 and 11 years old
Primary school	The name of the schooling system for children aged 4-11 years old within the UK
Key stage one	Legal term for the first two years of school for maintained schools throughout England and Wales, normally known as years 1 and 2, where children are aged between 5-7 years old
Key stage two	Legal term for the four years of school for maintained schools throughout England and Wales, normally known as years 3, 4, 5 and 6,, where children are aged between 7-11 years old

Chapter One

Introduction

1.1 Introduction

Physical activity (PA) is defined within the literature as "any bodily movement produced by skeletal muscles that results in energy expenditure" (Caspersen, Powell, & Christenson, 1985, p. 126). The numerous health, social and emotional benefits of habitual PA are also well documented. In particular, within child populations participation in regular PA can reduce the prevalence of cardiovascular disease (Ekelund, Luan, Sherar, & et al., 2012; Lang et al., 2018), while improving bone health (Nyestrom et al., 2018; Osborn et al., 2018; Baptista et al., 2012) and motor development (De Meeester, 2019; Laukkanen, Pesola, Havu, Sääkslahti, & Finni, 2014). Furthermore, children who are regularly active also experience increased levels of cognitive functioning (i.e. attention and academic performance) (Greeff et al., 2018), and improved social development, while typically reporting higher levels of self-esteem (Biddle, Ciaccioni, Thomas and Vergeer, 2018; Ahn, Sera, Cummins and Flouri, 2018). Notably, in order to achieve the outlined health benefits, children are encouraged to meet or exceed the current PA guidelines. Within the UK, Chief Medical Officer's recommends that children between the ages of 5-18 years participate in an average of at least 60 minutes of moderate-to-vigorous PA (MVPA) per day across the week and this guidance is similar globally (WHO, 2018; Chief Medical Officers' Physical Activity Guidelines, 2019). In addition, recent guidelines recommend that children and young people should engage in a variety of types and intensities of physical activity across the week to develop movement skills, muscular fitness, and bone strength (Chief Medical Officers' Physical Activity Guidelines, 2019).

Children's PA levels are a cause for concern within the UK and internationally, with statistics reporting that within the UK only one fifth of boys and 14% of girls (aged 5-16 years) are meeting the current guidelines (Public Health England, 2019). Similarly, within Canada only 35% of 5-17-year-olds are achieving the recommended activity levels for their

age group (participACTION, 2018). While in Australia the overall PA levels have been awarded a failing grade of D- (grades were assigned to a total of 12 indicators which were clustered under the categories: Strategies and Investments, Settings and Sources of Influence (Family and Peers, School, Community and the Built Environment), Overall PA Levels (Organised Sport and PA Participation, PA Participation in Schools, Active Play, Active Transportation, Screen Time) and Traits (Physical Fitness, Movement Skills. Grades for the report card indicators were informed using data synthesised from national based surveys and studies, collected from 2013 onwards. Overall PA levels were graded a D- and specifically within the 7-11 age range, device-measured (via accelerometery) data highlighted that on average 52% of primary school-aged children accumulated at least 60 minutes of daily MVPA) for the past 6 years, the national report card commenting that components of the PA guidelines such as muscle and bone strengthening activities have been 'forgotten' (Active Health Kids Australia, 2018). Interestingly, while 95% of children in England communicated that they enjoyed PA (Sport England, 2019) and this being a key factor for participation (Coulter, McGrane and Woods, 2019), activity levels remain low within the child population (Sport England's Active Lives Children and Young People Survey, 2018). As such, there is a need to explore alternative approaches to PA promotion as current practices are not encouraging the majority of children to be active.

Across the majority of Western countries, school attendance within the 7-11 age range is compulsory, thus making primary schools an optimal setting for PA promotion. Specifically, schools are considered to be nurturing environments where children have opportunities to be active, learn about PA and develop positive PA behaviours (Chróinínin, Murtagh and Bowles, 2012; Martin and Murtagh, 2015). However, it is reported that only 28% of children meet the guidelines of 30 minutes of MVPA per day within a school setting (O'Keefe, 2018). Further, while primary school physical education (PE) is considered to have

the potential to address many of the concerns raised about children's health, wellbeing and PA levels (Carse, Jess and Keay, 2017), it is reported that teachers are not provided with the tools necessary to support meaningful PE experiences (Chróinínin, Fletcher and O'Sullivan, 2018). The National Curriculum for Physical Education in England aims to ensure that all pupils develop competence to excel in a broad range of physical activities, are physically active for sustained periods of time, engage in competitive sports/activities and lead healthy, active lives (Department for Education, 2019). However, it is often reported that PE does provide ample opportunities for children to meet the PA guidelines (Smith, 2013; Keegan et al., 2015; Powell, 2014). Consequently, this suggests that in order to support regular PA within primary schools, the ongoing challenges within this setting require further exploration, particularly within the PE provision. Moreover, the perceptions of teachers, practitioners and children themselves are essential to guiding this process, ensuring that the primary school setting is adequately prepared for the implementation of PA practices.

In recent years, the concept of "physical literacy" (PL) has gained prominence in both research, policy and practice related to PA promotion in children. While there are various interpretations of the concept internationally, PL is typically defined as being "an individual's motivation, confidence, physical competence, knowledge and understanding to take responsibility for engaging in physical activities for life" (IPLA, 2017). As such, it has been proposed as a concept that challenges individual engagement in PA (Whitehead, 2001; 2010; 2019). Notably, Edwards et al.'s (2017) systematic review findings reported that the prevalence of PL related research publications increased steadily from 2001-2019. A marker of the recent popularity of the concept is the inclusion of PL within various international governmental policies. For example, ten countries internationally recognise PL either explicitly or implicitly in their policies and programmes thus highlighting the global uptake of the concept (Roetert and Macdonald, 2015; Shearer et al., 2018). Specifically, within the

UK, the recent School Sport and Activity plan (Departments of Education, Culture, Media and Sport and Social Care) recommended that all sport provision is based upon 'principles' of PL (Department of Education, 2019). Whilst, the Society for Health and Physical Educators in America released National Standards for PE stating that the incorporation of the term PL into practice is the ultimate goal of PE (SHAPE, 2014). However, while it is positive that PL is being supported on a policy level, there is a lack of practical advice or guidelines articulated for teachers and educational practitioners to implement, assess and foster PL practices within schools (Edwards et al., 2017; 2018; 2019; Giblin, Collins and Button, 2014).

PL assessment, in particular, has recently emerged as an aspect that requires further investigation (Barnett et al., 2019; Giblin, Collins and Button, 2014; Edwards et al., 2018). Similar to other core subjects within the school curriculum, without the ability to measure PL, practitioners, teachers and policy makers cannot ascertain if children are displaying positive PL behaviours or if additional support is required (Mandigo, Lodewyk and Tredway, 2019). In recent years' PL assessment tools have been developed such as the Canadian Assessment of PL (CAPL) (Longmuir et al., 2015) and the Passport for Life (PHE Canada, 2013). However, existing assessments have been criticised as being heavily weighted towards the physical domain (Robinson and Randall, 2016). Therefore, there is a substantial need to understand more comprehensively whether available tools are aligned to Whiteheads (2019) interpretation. Furthermore, it is pertinent to comprehensively explore existing tools to ascertain if they can be utilised to support the assessment of the concept or in the absence of such assessments, inform the direction of future tool development. Without an assessment of PL, it is impossible to track the progress of the concept among individuals and within schools on a local and international scale.

The evidence base within the discipline of PL is limited. However, recent research highlights a shift towards utilising established research methods to investigate the concept further. These include but are not limited to systematic reviews (Edwards et al., 2017; 2018) intervention studies (Cairney et al., 2019) and Delphi polls (Keegan et al , 2019). Recommendations for future research suggest that child populations should be the focus in order to promote positive trajectories for health (Utesch, Bardid, Büsch and Struass, 2019). In particular, the modified Strachan-Sheikh model (Strachan and, Sheikh, 2004) of life course health trajectories highlights that the 'early life', beginning at conception and ending with late adolescence is a crucial developmental period for health enhancing behaviours (Vineis et al.,2016). As such, the work outlined within this thesis predominantly considers the concept of physical literacy among children between the ages of 7-11 years, as this is thought to be an optimal age for developing learned behaviours (Steinbeck, 2001; Black et al., 2017).

1.2 Overview of the thesis

The work included within this thesis sought to explore PL assessment specifically within children aged 7-11 years with the aim of promoting and supporting lifelong engagement in PA. Chapter 2 of this thesis is a literature review that builds on the introduction chapter, providing a comprehensive review and critique of the current literature relating to PL, PE and assessment within the 7-11 age range. Key studies are included to inform discussions around the varying interpretations of PL, current PL practices and assessment of the concept. Specifically, the literature review highlights 'gaps' within the current knowledge base which have subsequently provided the rationale for the research contained in this thesis. In addition, Chapter 2 will also detail the aims and objectives of this thesis and review the methodological approaches that have been employed, providing detailed information on the mixed-methods approach utilised.

Chapter 3 presents Study 1, a review of international definitions and interpretations of the PL concept. Specifically, in Study 1, the global definitions of PL are investigated in order to collate and compare the varying interpretations of the concept. Study 1 aimed to provide clarity and inform further research into the concept on an international scale and as such, significantly influences the subsequent chapters within this thesis. Study 2 (Chapter 4) utilised a systematic review approach in order to identify and appraise existing assessment tools within the PL discipline and related research fields. Study 2 sought to not only identify existing assessments within the literature but also appraises the quality of each assessment with regards to the methodological rigour, feasibility and alignment to the PL concept. Study 2 has greatly informed the direction of this research programme, as the availability and quality of tools became apparent.

A crucial aspect of the tool development process is working alongside the target population to ensure the assessment is valid and appropriate for use within the required setting (Van Rossum et al., 2018). Therefore, Chapter 5 reports the findings of study 3, an investigation into stakeholder's perceptions of PL assessment. Study 3 embraced a qualitative approach and through participating in concurrent focus groups, teachers, children and self-defined PL experts were encouraged to share their experiences of assessment and give suggestions for a PL assessment that could be used within a primary school setting.

Chapter 6 aimed to collate the evidence reported within each of the aforementioned studies alongside utilising existing literature to produce informed recommendations to support the development of a novel assessment of PL within the 7-11 years population. The recommendations outlined within this chapter provide pragmatic solutions to the challenges faced with assessing a holistic concept underpinned by a strong philosophy. The purpose of the recommendations chapter is to enable practitioners, teachers and children to start assessing PL, thus encouraging reflection and progression within their unique PL journey.

Finally, Chapter 7 provides a synthesis of the work included within this thesis. The strengths and limitations of each study are discussed in detail alongside the originality and novel approaches that have been utilised within this project of work. In addition, numerous recommendations are included within this chapter to guide the direction of future research. These are based predominantly on the findings from the studies included within this thesis and the knowledge gained throughout the duration of the project of work.

1.3 Thesis Study Map

For the benefit of the reader each of the three study and recommendations chapters in this thesis will be prefaced by a thesis studies map, outlining the objectives and key findings of each study. The aim in presenting the 'map' is to efficiently summarise the key objectives and findings of the studies and 'map' where each study fits within the overall thesis. In addition, the flow of studies is presented in Figure 1 and highlights the direction of information.

Figure 1.1 Flow of information throughout the PhD programme of research

- Global interpretations of physical literacy

2017

2017-2018

- Systematic review of existing assessments

- Qualitative exploration of stakeholders perceptions

2018-2019

Recommendations for the development of a physical literacy assessment tool for use within children aged 7-11 years
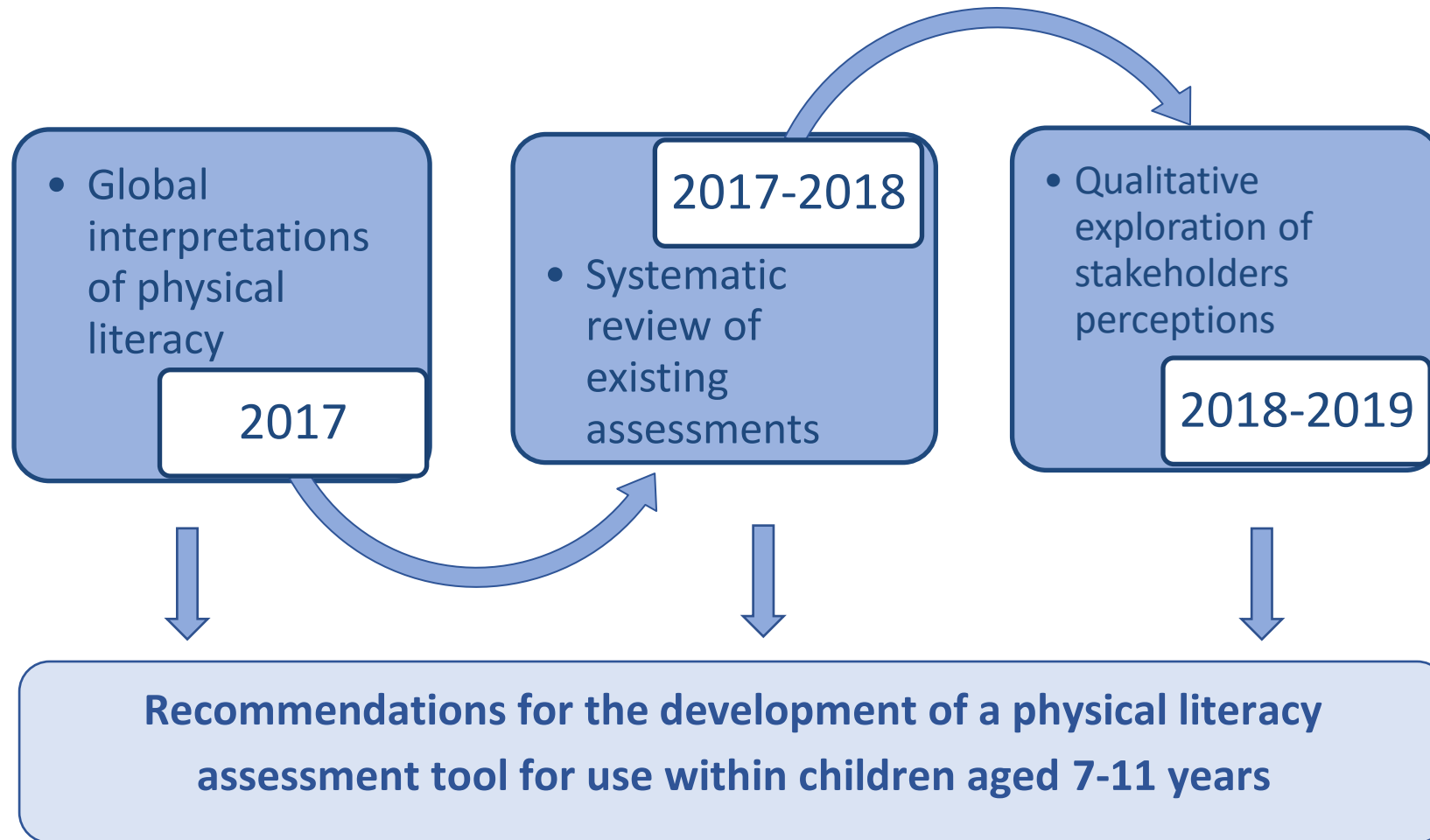
Chapter Two

Literature Review

2.1 Introduction

The purpose of this chapter is to review the literature associated with the concept of PL, specifically within the child population aged 7-11 years. This literature review will seek to: (i) define and discuss the PL concept, (ii) review and critique the associated research in this field conducted to date, providing a clear rationale for this thesis and (iii) consider the demands and challenges associated with the assessment of PL and PE within the 7-11 years old age range. Finally, this chapter will conclude with the aims and objectives of the thesis and a justification of the methodological approaches that have been used within it.

2.2 Physical Literacy

The concept of PL has gained considerable interest and momentum in recent years due to the global physical inactivity crisis prompting further investigation of current PA practices (Bauman et. al, 2018; Cairney et al., 2019). As low levels of PA participation cause concern, researchers, practitioners and the government (i.e. Sport England) continue to strive to find a means of engaging the population in lifelong participation in PA with the aim of improving health and well-being (De Meeester, 2019; Lang et al., 2018; Biddle, Ciaccioni, Thomas and Vergeer, 2018). As such, PL has been proposed as a concept that challenges individual engagement in and promotion of PA (Whitehead, 2001; 2010; 2019).

2.3 Defining Physical Literacy

PL is a multidimensional concept that in the past decade has ignited debate within the health and education sectors through the work of education specialist Margaret Whitehead (Whitehead, 2001, 2007, 2010, 2019). Whitehead has been referred to as a "modern day champion of PL" (Cairney et al., 2019) and defines the concept as "the motivation, confidence, physical competence, knowledge and understanding to value and participate in

PA throughout the life course" (International Physical Literacy Association, 2017). In comparison to sporting or fitness-based approaches, PL is considered to be an alternative approach of encouraging PA participation due to its holistic nature that is underpinned by the philosophical ideologies of monism, existentialism and phenomenology (Durden-Myers, Whitehead, and Pot, 2018).

Contemporary research exploring PL has provided a diverse array of perspectives (Dudley, Cairney, Wainwright, Kriellaars, & Mitchell, 2017; Edwards et al., 2017; Keegan et al., 2019; Barnett et al., 2019). Most recently, the PL elements of motivation, confidence, physical competence, knowledge and understanding have been categorised into cognitive, physical, and affective domains (Whitehead, 2019). According to Whitehead, the physical domain includes the consideration of an individual's movement vocabulary (manipulation e.g. grasping), movement capacities (simple (balance), combined (agility) and complex movements (hand-eye coordination)) and movement patterns (general and refined). Specifically, Whitehead (2010, pg. 44) refers to a hypothetical individual who is 'physically competent' as being able to "move with poise, economy and confidence in a wide variety of physically challenging situations" while elaborating that this is inclusive but not limited to body management, coordination and control. While Whitehead (2010; 2019) discusses various attributes of physical competence, the element itself is not precisely defined, thus, making it challenging to identify appropriate means of assessment.

The affective domain of PL is related to an individual's confidence and motivation to participate in regular PA, however, the sub-elements within the domain have not been identified (Whitehead, 2010; 2019). Whitehead's recent work proposes that an individual's motivation includes the desire to participate in PA, engaging in challenging activities and, ultimately, committing to PA as a part of daily life (Whitehead, 2019). In addition, the confidence to participate in PA compliments aspects of an individual's motivation within the

affective domain, and as such, the two are often considered in tandem (Whitehead, 2010; 2019). Notably, an individual who exemplifies 'maximising potential' within their PL journey has confidence in their ability to enhance current knowledge, achieve predetermined targets/goals and maintain regular practice or effort (Whitehead, 2019). Furthermore, while the PL domains (affective, cognitive and physical) are communicated individually, the elements within each domain (i.e. affective; motivation and confidence, physical; physical competence and cognitive; knowledge and understanding) are often considered together and this is encouraged due to the holistic nature of the concept.

The cognitive domain is a further facet of PL, with Whitehead articulating the domain as "the knowledge and understanding that is required to engage in an active lifestyle throughout an individual's lifetime" (Whitehead, 2010). Specifically, the attributes of the cognitive domain relate to an individual's ability to comprehend, retain and utilise relevant knowledge that is specific to their individual PL journey (Whitehead, 2019). This includes the ability to analyse aspects of movement, reflect/inform judgements on progress and maintain a sound understanding of the lifelong benefits of participating in habitual PA (Whitehead, 2019). Recent research from Cale and Harris (2018) further explores the cognitive domain, providing pedagogical approaches and practical strategies to develop and assess knowledge and understanding, specifically within the key stage 2 population (7-11 years). It is suggested that children within the 7-11 age range should be able to explain the need for safety/rules, explain and feel benefits of exercise, explain some physiological benefits of exercise and identity opportunities to be active while also explaining that being active is a personal experience. Further, it is suggested that alternative approaches for both PA promotion and assessment should be explored to understand pedagogy that aims to foster PL (Cale and Harris, 2018). In particular, Whitehead's recent work proposes that individuals should be encouraged to gather 'data' relevant to their own PL journey, supporting the

process of making informed, impactful and individualised decisions. This is in line with previous research that demonstrates that self-monitoring is a well-documented behaviour change technique within the PA research discipline (Sallis and Saelens, 2000; Hayes and Van Camp, 2015; Hansen et al., 2019). However, with limited PL assessments available, it is currently unclear how self-monitoring would be supported in practice (Edwards et al., 2018).

As discussed above, Whitehead has articulated the overarching domains of PL an provided descriptive examples of PL attributes. However, other interpretations and definitions of the concept are generally unknown. In addition, it is important to consider the recognised overlap between Whitehead's articulation of PL and the terminology used within various well-established research fields (Keegan et al., 2019). For example, in the field of motor competence the importance of developing object control, locomotor skills, balance and are currently well established as precursors to participation in PA (Hulteen et al., 2018). As such, various assessments are currently available for use within the 7-11 age range (Barnett et al., 2019; Barnett et al., 2016; Rudd, Butson and Barnett, 2016). Thus, within the PL research discipline there is a substantial need to further explore the links with related research fields, alongside different interpretations both nationally and globally of the concept. In doing so, relevant existing assessments and resources could be utilised to support the assessment of the concept. Furthermore, evidence-based research into the concept is still in its infancy; therefore, it is arguably only natural that varying interpretations will emerge as researchers, practitioners and governmental agencies strive to continue to understand PL. Therefore, in order to facilitate appropriate assessment of the concept differing definitions and interpretations must be investigated with the purpose of guiding and supporting the implementation and assessment of PL.

2.4 Underpinning philosophy

Exploration of the history of PL suggests that while the concept may have originated in the early 19th century, the philosophical lens in which Whitehead views the concept is a novel perspective (Cairney at al., 2019). According to Whitehead, PL is underpinned by the three philosophies of monism, existentialism and phenomenology (Whitehead, 2007, 2010, 2019). Specifically, the PL philosophy of *monism* establishes that the mind and body are equal partners working in unison rather than separate entities, thus contradicting traditional dualistic approaches to PA promotion (Pot et al., 2018). Additionally, the existentialist aspect of the concept encourages the consideration of an individual as such, respecting and acknowledging that each embodied experience is unique and has the potential to shape and, ultimately influence interactions with the environment (Whitehead, 2010). Specifically, the richer and more varied an individual's PA experiences are, the more opportunity said individual has to flourish and realise their own unique potential (Pot et al., 2018; Merleau-Ponty, 1968). Finally, the phenomenological aspect of the philosophy considers the complex interactions humans have with their surrounding environment and how this can affect their relationship with PA. Whitehead promotes the holistic and inclusive nature of the concept by articulating monism, existentialism and phenomenology throughout her various works (2001; 2007; 2010; 2019), advocating that they should be interwoven to form the philosophical foundation of PL.

Whitehead's intention (cf. Whitehead, 2010), by invoking a philosophical stance, was to transform PL into a concept focused on the individual in the world and her/his experiences. However, the philosophical nature of the concept has ignited considerable debate within both the education and health sectors (Robinson and Randall, 2017; Tremblay and Lloyd, 2017; Corbin, 2016). The ongoing discussion surrounds the ability to implement, assess and operationalise PL while also aligning with the underpinning philosophies (Giblin, Collins and

Button, 2014; Whitehead et al., 2018; Edwards et al., 2018). Whitehead (2010) argued that an individual cannot fully understand or appreciate the true nature of PL without first grasping its philosophical roots. Yet for many, the detailed and complex philosophical groundings of PL present a barrier to clarity and understanding (Jurbala, 2015). Moreover, with regards to assessment, within a systematic review of existing measurement tools, authors have questioned the likelihood of an individual's ability to assess PL progress without a sound understanding of the philosophy and ontology (Edwards et al., 2018). The review by Edwards et al. (2018) then continues to report that 33% of included studies did not acknowledge any philosophical considerations within the assessment process. Potential explanations for this include a lack of transparency or value being placed on the philosophical aspects within existing assessment tools and that the included assessments were not developed specifically for PL (Edwards et al., 2018). Furthermore, a recent Delphi poll including PL experts (n=18) with the aim of providing agreement within the concept has positioned the philosophy as a core aspect of PL and suggested that bypassing the philosophical aspects due to lack of understanding would be counterproductive (Keegan et al., 2019). Therefore, regardless of the complexities, it is crucial for researchers and practitioners to consider the philosophical aspects. Further, this is of particular importance in regard to methods of assessment as the holistic approach offered by PL is significantly different to the "one size fits all" approaches commonly illustrated within the education and health sectors (Pot, Whitehead and Durden-Myers, 2018).

2.5 Evidence for Physical Literacy

Whitehead argues that PL is a concept that should be encouraged throughout all stages of life; however, a prominent concern within the PL field is the lack of empirical evidence available to support the concept (Cairney at al., 2019; Edwards et al., 2019; Shearer

et al., 2018). Instead, what is common within the discipline is 'academic opinionating', which has been criticised as leading to misconceptions surrounding the concept (Hyndman and Pill, 2018). However, contemporary literature has provided a conceptual framework based on existing empirical evidence hypothetically linking PL to physical, social and mental health (Cairney et al., 2019). Notably, while PL is often considered to be linked to health, no prior evidence has been presented to support this, potentially due to the lack of PL assessments available for use in both research and practice. The theoretical model proposes a bi-directional relationship with regular PA participation and physiological, social and psychological adaptations which are then linked to the prevalence of physical, mental and social health. Cairney et al. (2019) proposed that by viewing physical literacy as a determinant of health, this would provide conceptual direction for empirical physical literacy research. However, whilst this is one of the first steps to highlight the role physical literacy may play in promoting health behaviours, the model itself draws heavily from existing evidence from fields outside of physical education, such as motor competence research (Stodden et al., 2008). As discussed previously within the introduction chapter of this thesis, the links between PA and health outcomes are already well established within the literature. However, there is a lack of clarity as to whether it is PL or PA that provides health benefits within the presented model. Specifically, there is a need to determine whether PL elements (i.e. motivation, confidence, physical competence, knowledge and understanding) empirically lead to improved health and wellbeing.

While evidence-based research is limited within the PL discipline, recent offerings highlight a shift towards utilising established research methods to investigate the concept further. These include but are not limited to systematic reviews (Edwards et al., 2017; 2018), intervention studies (Cairney et al., 2019; Sheenan et al., 2010), surveillance studies (Lang et al., 2018; Silvia et al., 2018) and Delphi polls (Keegan et al., 2019; Shortt et al., 2019;

Francis et al., 2016). Moreover, exploratory research investigating the correlations between

PL domains and any links to fitness and weight status have started to build the much needed

evidence base (Nyström et al., 2018). However, there is a clear need for more empirical

evidence that investigates PL in children to understand, in greater depth, what is needed to

support children to be active for life. In order to investigate and understand the concept of PL

within the child population it is important to determine the best assessment methods to

generate evidence that has the potential to inform both policy and practice.

2.6 Physical Education

An important context for the promotion of PL is Physical Education (PE). The aim of primary

school PE for children aged 5-11 years old is to encourage a positive relationship with PA

through improving children's confidence, motivation and social skills while facilitating the

ongoing physical development of each child (Department of Education, 2017). The

conceptualisation of PL continues to challenge traditional PE practices, thus encouraging PE

to develop from lessons based predominantly on competitive sports and skill acquisition

towards a holistic approach centred around the overall health and development of the child

(Jurbula, 2015; Giblin, Collins and Button., 2014). As stated in the introduction section,

recently, within the UK, PL has been included within governmental policy to support the

development of PA provision in schools. The Sporting Futures document produced by the

Department of Education included 'increase in the percentage of children achieving PL

standards' as a key performance indicator (Department of Education, 2015). Similarly, in the

recent School Sport and Activity plan (Department of Education, Department of Culture,

Media and Sport and Department of Social Care, 2019) it is recommended that all sport

provision is based upon 'principles' of PL (Department of Education, 2019). While it is

positive that PL is being considered on a policy level, there is a lack of practical advice or

guidelines articulated for teachers and educational practitioners to assess PL within schools. As such, feasible and psychometrically sound assessment tools are required to ascertain where children are on their PL journey.

Whitehead (2001) expresses that it is not the sole responsibility of PE to implement PL practices within the curriculum. However, due to the practical nature of the subject area and the range of facilities typically available within a PE setting, it is still deemed to be an appropriate starting point for operationalising the concept within an education setting (Lundvall, 2015; Edwards et al., 2016). Within the UK, The Department of Media Sport and Culture (DCMS) provided ten outcomes, which might be observed if children were experiencing high quality PE (DCMS, 2010). Outcome (1) 'a commitment to making PE/sport a part of their lives within and out with a school context' aligns with the purpose of PL 'to value and engage in PA throughout the lifetime' (IPLA, 2017). Additionally, there are also links with the cognitive domain in outcomes (2) 'know and understand what they need to achieve and how to do so', (3) 'understand that PE and sport are an important part of a healthy active lifestyle', and (7) 'think about what they are doing and make appropriate decisions for themselves' (Department of Media Culture and Sport, 2010).

Notably, whilst the literature surrounding PL and high-quality PE highlights the similarities of the concepts, it also outlines the differences (i.e. the sport specific nature of high-quality PE compared to the PA focus of PL) and stresses that these terms should not be used interchangeably (Lounsberry and McKenzie, 2015). Specifically, Whitehead refers to high quality PE and PL as 'discrete concepts' that are not in competition, instead, both essential in their own right (Whitehead, 2019, pg. 28). It is not surprising, however, that at times the terms are confused as the PL concept aligns closely with the aspects of high-quality PE in schools (Department of Education, 2015; 2019). Nevertheless, PL is viewed as an approach that could change the way in which PE is delivered, assessed and accepted by

teachers, children and practitioners (Robinson and Randall, 2018). Much like PL, primary school PE is considered to have the potential to address concerns surrounding children's health, wellbeing, sport participation and PA levels (Jess, Keay and Carse, 2016). Therefore, PL assessment may be able to support and improve the implementation of PE within a primary school context as both variables undoubtedly strive to support the positive physical, cognitive and affective development of children.

2.7 Assessment of Physical Literacy

Assessment, in its most rudimentary form, refers to the collection of information (Smith, 1997). Within the education sector, assessment is thought to be essential for charting an individual's progress and highlighting key areas for development both on an individual and population level (Hay and Penny, 2012). Traditionally, assessment in PE, and arguably the education sector as a whole, has placed an importance on summative assessment (formal assessments to evaluate learning at the end of a topic or unit of work), e.g., Standard Assessment Tests (SATs), General Certificate of Secondary Education (GCSE), Secondary School Literacy Test (Department of Education, 2017; Cheng, Klinger and Zheng, 2007). While formative assessment, defined as "activities undertaken by teachers— and by their students in assessing themselves—that provide information to be used as feedback to modify teaching and learning activities" (Black & Wiliam, 2010, p. 82) is often lacking within both policy and practice. Consequently, education practitioners in recent years have stressed the need for both formative and summative assessments within a primary school context, while ensuring that these are based on assessment-for-learning approaches (Stiggins, 2002; Dixon and Worrel, 2016).

Assessment for learning, also known as formative assessment, refers to any assessment or testing activity that informs the learning process (Williams, 2011). Unlike summative

assessment which evaluates pupil knowledge and achievement after a period of learning is complete, assessment for learning involves evaluating small content areas as part of the ongoing learning journey. While summative assessment is essential for evaluating learning over a specific time period, assessment for learning approaches can encourage the assessor to utilise a variety of different tools and this process is often student led (Dixson and Worrell, 2016). For example, within the UK, the National Foundation for Educational Research report details self-assessment and peer assessment as examples of assessment for learning, classifying these as 'informal' assessment styles (NFER, 2007). The purpose of assessment for learning is to create a culture where children can make mistakes and learn through receiving descriptive feedback via written comments or verbal conversations. By providing descriptive feedback and co-creating learning goals, progress can be shared with parents or carers outside of the classroom (Black & Wiliam, 2010).

According to Hay and Penney (2009), assessment within PE should be considered as a process through which learning can be promoted. However, within the UK, guidance for assessment in PE is not included in the national curriculum. As it stands, there is no national approach to assessment in PE unlike other subjects such as maths and English (Department of Education, 2019). Ultimately, this makes it increasingly difficult for teachers and educational practitioners to chart progress of PL and high-quality PE outcomes. While there have been various assessments created by sporting organisations (e.g., TOP Start developed by the youth sport trust) or PE specialists (e.g., PE passport https://www.primarypepassport.co.uk/) these typically only provide assessment of elements within the physical domain of PL and not the affective and cognitive domains. Further, little is known about the acceptability and implementation of assessments in PE particularly within practitioner groups (i.e., teachers, coaches, and classroom assistants) and this information is vital to understanding if a tool can be used within the primary school context.

Attempting to assess a complex and interchangeable concept such as PL is a challenge (Edwards et al., 2017; Giblin, Collins and Button, 2014). However, within the education sector assessment is a foci on which curricula are informed and standards measured, therefore to give PL longevity, types of assessment must be considered - perhaps not in a traditional sense but with an assessment-for-learning approach being used as a frame of reference (Giblin, Collins and Button., 2014; Corbin, 2016). Furthermore, assessment is crucial when considering an individual's PL journey; for example, is a child's relationship with PA changing and if so, is it improving or declining? In order to chart progress on an individual and population level, appropriate measurement of PL is required with the results of such assessment informing best practice within the education sector. Despite PL being a lifelong journey, given the importance of the childhood years for establishing PA behaviours, assessments within the 7-11 age range could be utilised within the education sector to support the development of PL and inform best practice. As such, Whitehead (2019) has recently responded to the claims that the philosophy is difficult to incorporate within assessment and practice by producing recommendations that align with the philosophy for charting an individual's PL journey (see **Table 2.1**).

Table 2.1 Relationship between philosophical principles underpinning PL and assessment (Whitehead, 2019, pg.76).

Philosophy	*Implications for charting a PL journey*
Monism	Information captured from across all three domains, affective, physical and cognitive. All procedures designed to promote motivation.
Existentialism	Information captured in a wide variety of situations, including those involving different environments, different PA protocols and different relationships to others. Judgements regarding constituent aspects of participation (e.g., techniques) a relatively small consideration.
Phenomenology	Individuals treated as unique. No comparisons made with others. Judgements ipsative (set against personal data).

While Whitehead (2007; 2010; 2019) has consistently advocated for the need to consider the philosophy within PL assessment, this is the first step towards providing pragmatic guidance to enable teachers, practitioners and researchers to do so. Additionally, providing examples of the philosophy in practice also potentially facilitates an understanding of the philosophy as a whole. While this is a positive direction for PL assessment, there is also a need to consider how each philosophical aspect may already be represented within existing assessments.

A recent review from Edwards et al. (2018) identified 52 assessments of PL and related constructs, evaluating these in relation to age group, environment, and philosophy (Edwards et al., 2018). The assessments that reported an explicit distinction regarding alignment to a physical literacy domain were categorised further as sub themes: 22 tools

examined the physical domain, eight the affective domain and five assessed the cognitive domain. Notably, only one assessment provided a measure of all three physical literacy domains (physical, cognitive and affective) within the tool, while 66% of assessments were used in children under 12, suggesting this age group as being a key area in existing research (Edwards et al., 2018). However, while 52 assessments were identified within the review, Edwards and colleagues may have missed some relevant assessments as their search terms were limited to 'PL' and did not include search terms specific to each element of PL (i.e., motivation, confidence, physical competence, knowledge and understanding). In addition, while alignment to the philosophy was appraised within the review, the precise methodological process of doing so was unclear. Subsequently, leaving uncertainty regarding the alignment of each tool with Whiteheads' articulation of the concept and prevents examples of best practice being highlighted. Moreover, the review did not examine the psychometric properties of identified assessments despite validity and reliability of PL assessments been criticised within the literature (Longmuir and Tremblay, 2016). Further, while the assessments were categorised under the lower order themes of physical, cognitive and affective, specific appraisal information regarding the psychometric properties, feasibility and alignment of each tool to the concept was not reported. Key information and detail is required for existing assessment tools related to PL, such as what elements are being assessed within each domain, the psychometric properties of each tool and the feasibility of using each tool within a primary school setting.

In recent years' PL assessment tools have been developed such as the Canadian Assessment of PL (CAPL) (Longmuir et al., 2015) and the Passport for Life (PHE Canada, 2013). Whilst both assessments are continually undergoing redevelopment, it is thought that these are the first points for the assessment of PL in child population (Tremblay et al., 2017). However, both CAPL and Passport for Life have been criticised within the literature due to

their misalignment to Whitehead's conceptualisation and being heavily weighted towards the physical domain (Robinson and Randall, 2016). For example, the most recent version of the assessment CAPL-2 gives children the opportunity to score 30 points for each of the physical domain, affective and added behavioural domain, while only 10 points within the cognitive domains, thus affecting their overall PL 'score'. Moreover, research investigating correlations between PL domains reported that in terms of the physical domain there was a moderate correlation with the affective domain and a weak correlation with the cognitive domain (Nyström et al., 2018). Therefore, it may be pertinent to link indicators within the affective/physical domain within an assessment but not necessarily the cognitive domain. There is a need to understand the relationship between the domains further in order to best inform PL assessment. Whilst both CAPL-2 and Passport for Life represent the first available tools for PL assessment within a primary school population (Tremblay and Lloyd, 2010; Tremblay, 2017), there is a considerable need to understand the relationship between domains and the philosophical alignment of existing assessments to ensure that the concept is being assessed effectively (Corbin, 2016; Lundvall, 2015).

Contemporary research has provided a user guide for selecting suitable assessments of PL in an attempt to encourage research and practitioners to utilise existing assessment tools where possible (Barnett et al., 2019). As previously discussed, despite being a concept that has only recently been considered within the education and health sectors, many of the PL elements may be represented in assessment across various psychology, sports science and cognitive development fields (Barnett et al., 2016; Hulteen et al., 2018; Rudd, Butson and Barnett, 2016). The user guide provides steps for selecting an appropriate tool and discusses considerations for tool selection. This is in contrast to many of the debates regarding definition and interpretation - instead of having one assessment there can be many depending on the context or scenario. The user guide provides the assessor with seven different steps to

consider before selecting 'the most suitable method'. The initial steps involve prompting the

assessor to consider the elements to be assessed, context of assessment and target population.

The final steps consider the feasibility of such methods, i.e. number of participants and cost

efficiency. This is potentially an interesting approach for teachers, researchers and

practitioners as it encourages the breakdown of each PL domain into manageable and

assessable parts as opposed to one complex and time-consuming assessment. For example, a

teacher may not have the budget or expertise to assess the entire PL of a child at one time

point so they could alternatively select the physical domain and assess one element of that

domain, e.g. agility. However, while this user guide is a pragmatic approach aimed at guiding

the assessment of PL, further information on existing aligned assessments, the psychometric

properties of each tool and the feasibility of assessments within a school setting is needed.

2.9 Feasibility
Considering the feasibility of an assessment tool is of vital importance when determining if

the assessment is appropriate for the use within an educational context (Barnett, Dudley,

Telford and Lubans, 2019). However, despite this, the feasibility of a tool is often not

reported or overlooked within the developmental stages (Klingberg, Schranz, Barnett, Booth

and Ferrar, 2019). When considering PL assessment within a primary school setting, it is

necessary to take into account the specific contextual requirements (Corbin, 2016; Lundvall,

2015). Specifically, teachers have listed barriers such as lack of time, space and expertise as

barriers to conducting physical assessments within a primary school setting (Lander et al.,

2016). Within the literature, it is often considered that expert opinion is crucial to informing

the developmental process of an assessment tool, particularly at the design phase when

appraising the feasibility of an assessment (Longmuir, et al., 2018). However, within a

primary school setting, it is the teachers, support staff and children who have the expertise on

what would be feasible and acceptable within this context (Morley, Van Rossum, Richardson and Foweather, 2019).

While there is a plethora of literature available relating to barriers of assessment practices for teachers, within a school context there is a gap in the literature with regards to the acceptability of assessments. Specifically, while PL tools are available for use within a primary school setting (e.g., CAPL-2, passport for life), the feasibility of these have not been reported consistently (Lundvall, 2015). Furthermore, in general research-grade tools have typically not been accepted within a primary school context, due to being too lengthy, complex and requiring expertise to administer or feedback results that a typical classroom teacher or practitioner may not possess (Lander, Hanna, et al., 2017; Van Rossum, et al., 2019). An investigation into the feasibility of existing assessment tools and ascertaining the perspectives of key stakeholders is essential to considering the reality of administering PL assessments within a primary school setting.

2.10 Measurement properties

Measurement properties are essential components to consider with regard to developing an assessment tool (Terwee et al., 2007; Monkkink et al., 2018). For a tool to be deemed appropriate for use, the validity (the accuracy of a tool to assess the construct it is claiming to assess), and reliability (the consistency of an assessment tool) must be established within the populations of interest, this is typically demonstrated using psychometric and theoretical testing (Robertson et al., 2017). The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) consider a 'gold standard' approach for reporting psychometric properties to include information relating to the testing of: Patient-Reported Outcome Measures (PROMs) development, content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing and responsiveness (Terwee et al., 2017). Recent updates from COSMIN

suggest that the PROM development and content validity are among the most important

measurement properties, proposing that if a tool is not proved to be valid for use within the

targeted population then arguably, they are not appropriate (Mokkink et al., 2018; Welk,

Corbin and Dale, 2000; Guyatt, Walter and Norman, 1987). Specifically, COSMIN

encourages the following components to be considered in terms of PROM development and

content validity: Relevance (i.e. Are the included items relevant for the construct of

interest?), Comprehensiveness (i.e. Are key concepts missing?) and Comprehensibility (i.e.

Are the PROM instructions understood by the population of interest as intended?) (Mokkink

et al., 2018). With particular regard to the comprehensibility aspect of PROM development,

this further supports the need to include the target population in the developmental stages of

an assessment tool. Specifically, by working closely with tool users to investigate if included

items are relevant for the population/construct best practice can be established in regard to

content validity and PROM development (Mokkink et al., 2018; Terwee et al., 2019).

Additionally, while the COSMIN guidance is specific to health and medical grade

assessment tools, Robertson et al. (2017) have outlined the properties that are essential to a

sport and exercise science context. Through reporting the findings from a two-round Delphi

poll (n=33) including exercise scientists, academics and sport scientists, consensus (67%

agreement) was reached on the measurement properties included in Table 1.2. While

properties such as reliability are defined consistently, within a sport and exercise science

context properties such as floor and ceiling effects are also deemed to be important

(Robertson et al., 2017; Goodway, Ozmun and Gallahue., 2019).

Table 2.2 Measurement properties, adapted from (Robertson et al., 2017)

	Level 1	Level 2
Reliability	Re-test reliability Intra-rater reliability Inter-rater reliability	Stability Internal consistency
Validity	Content Discriminant	Convergent Concurrent Predictive
Responsiveness	Responsiveness/sensitivity Minimum important difference/ smallest worthwhile change	Floor and ceiling effects
Feasibility	Interpretability Familiarity required Duration	Scoring complexity Completion complexity Cost

Within the PL assessment literature, a select few psychometric properties of available tools have been considered, however, the direction of research within this area is highlighting that it is essential to report a full range of measurement properties in order to inform appropriate tool selection (Mokkink et al., 2018; Robertson et al., 2017). Furthermore, previous reviews have not reported psychometric properties in detail (Edward et al., 2018). As such, it is important to extensively search the literature to identify and collate any of this information related to the psychometric properties of existing assessment of PL to establish a 'full picture' of methodological rigour. This will enable well-informed tool selection, which in turn will allow researchers, teachers and practitioners to have confidence in the results of PL assessments.

2.11 Physical Literacy Alignment

Despite many of the PL elements such as confidence, motivation and physical competence being well established within other research fields (Gallahue, Ozmun, & Goodway, 2012; Ulrich, 2000; Lubans et al., 2010; Hulteen et al., 2018), little is known about existing assessments and their alignment to the PL concept. While an ideal assessment would align both to the concept and the underpinning philosophy, Whitehead's (2019) recent work highlights that it may be the assessment experience itself that lends to the philosophy, i.e. Phenomenology- individuals treated as unique, no comparisons made with others. Therefore, there is a need to consider the content of existing assessments, if they assess the elements of PL and if they can be administered in line with the philosophy. While some of the PL elements are well known and assessments well cited within the literature, i.e., physical competence and fundamental movement skills (Hulteen et al., 2018), less is known about the other elements or elements within different contexts, for example, movement specific to an environment (Araujo & Davids., 2009; Renshaw, Chow, Davids & Hammond., 2010). Additionally, while the elements may be represented individually within existing literature, there is a lack of information regarding the assessment of the concept as a whole (Edwards et al., 2018). This holistic approach to assessment is prominent throughout Whitehead's various works (2007; 2010; 2019). Therefore, collating a descriptive overview of all existing assessments and reporting their alignment to the concept in a transparent manner will not only inform tool selection (Barnett et al., 2019) but also direct researchers to areas where assessments are lacking.

2.12 Summary

PL is a concept that has the potential to influence the way in which individuals understand, provide support and assess an individual's relationship with PA. The goal of PL is to promote participation in PA throughout the life course and for an increasingly inactive population.

Identifying or developing a feasible, psychometrically sound and aligned assessment tool is crucial to charting progress (Tremblay et al., 2016). Furthermore, a prominent concern within the PL field and one that has been discussed throughout this review is the lack of research evidence available (Edwards et al., 2019). Therefore, it is important that future research in this area considers: i) the various interpretations and definitions of the PL being implemented on a national and international scale, ii) existing assessments that may have the potential to support PL assessment across related research fields within the physical, cognitive and affective domains, iii) stakeholder perceptions of physical literacy and assessment practices within the primary PE provision, iv) produce evidence based, robust research that investigates the PL concept in its entirety (physical, cognitive and affective domains), and v) the development of initiatives to assess, promote and improve PL practices among young children between the ages of 7-11 years

2.13 Research aims and objectives

This thesis will seek to discuss the concept of PL, in particular exploring assessment. The overall aim of the work included within this thesis is to provide pragmatic recommendations to inform the assessment of PL within children aged 7-11 years within a primary school setting. As such, this aim will be achieved throughout the following objectives within each chapter.

Study 1: Global interpretations and definitions of Physical Literacy

- To collate, compare, and critically review existing definitions of PL from leading organisations implementing PL agendas around the world

Study 2: A systematic review of assessments related to Physical Literacy among children aged 7-11 years

- To systematically review the academic literature for tools to assess the physical, affective and cognitive domains of PL within children aged 7-11 years.

- To explore and critically discuss each assessment tool within the physical, affective and cognitive domains to appraise its (a) psychometric properties; (b) feasibility for use within a primary school setting and (c) alignment to the PL concept.

Study 3: Stakeholder perceptions of a Physical Literacy assessment for children aged 7-11 years

- To explore key stakeholders (academics/practitioners, teachers, and children) views of current practice, future directions and effective implementation of PL assessment, through concurrent focus groups

- To inform the development of a rigorous, authentic, and feasible PL assessment for use with children aged 7-11

2.14   Researcher positionality

In order to understand my positioning within the research, it is important to firstly outline my researcher background and philosophical stance. I view myself as a pragmatic researcher; I believe that there are various lenses in which to view and interpret the world when undertaking research, and that no single stance can give a complete understanding of an individual's experiences (Poucher, Tamminen, Caron, & Sweet, 2019). As a pragmatist, I link the choice of methodological approach directly to the purpose of and the nature of the research questions posed (Creswell & Creswell, 2018). To ensure the methodological coherence of a study, throughout this thesis I have explained in detail the rationale for each approach to ensure that the selected methodologies align with each research question. I fully, acknowledge that individuals create their own realities, based on their individual interaction with the world around them (Lincoln, 1995; Ponterotto, 2005), but also acknowledge my role in co-creating their realities. As such, the interpretations made throughout this PhD may have

been influenced by my own understandings and experiences, especially within the context of PA and PL. It is therefore important to share relevant aspects of my past with the reader, and to acknowledge that the interpretations presented in this PhD should be viewed as one of potentially many perspectives. Academically, I have completed a degree in Physiology and Sport Science, and it was through my honours research project that I first encountered the term PL. As the honours project focused on the PE provision within primary schools at the time my knowledge and understanding of the PL concept was limited. Throughout my PhD project, my comprehension and personal connection to the PL concept has increased considerably. As such, the direction of the research included within this thesis is a reflection of changes in my understanding of the concept over time, in particular opening my mind to connect with the philosophy on a deeper level and truly embrace the holistic nature of the concept. On a personal level, I have always perceived myself to be a physically active person, but the project has allowed me to reflect on my own relationship with PA and the progress I am making within my own PL journey. Due to the fact that I have persistently valued and engaged in PA throughout my lifetime, my beliefs and experiences of PA may have inadvertently influenced the participants within this project. This is something that should be taken into account when considering the findings within this thesis, particularly when I am representing stakeholder's experiences of PA, PL and assessment through qualitative means.

2.15 Independent Contribution to the Thesis

The work included within this thesis is part of a substantial and wider research project and as such it is necessary to be transparent with regard to the research conducted by me personally. Personal characteristics of individuals who supported studies included in Appendix 1C. Throughout the project I worked closely with a colleague (HG), to share, interpret findings and discuss the direction of research within the project. HG is investigating PL assessment

within children aged 5-7 years while my research focuses predominantly on children aged 7-11 years. For transparency my independent contribution to the primary research included in this thesis is documented below:

Study 1: Independent searches of the literature were conducted (CS, HG) to review and collate the various definitions and interpretations of PL internationally. After searching the literature, results were discussed at length (CS, HG) and divided the international groups to explore each interpretation further; Physical and Public Health Canada (PHE) (CS), Canadian Sport for Life (CS4L) (CS), Sport New Zealand (CS), SHAPE America (HG), IPLA (HG). For publication, writing the introduction section (CS, HG, EDM, RK), results/discussion sections relevant to each country IPLA (EDM), Australia (RK), Wales (LE), Canada (CS), New Zealand (HG). Preparing the manuscript for publication (CS, HG).

Study 2: Developing a search strand (CS, HG, LF, EDM). Conducted independent searches of the relevant electronic databases and the various screening rounds independently, improving the reliability of results (CS, HG). Data collection and quality appraisal; physical domain (CS), affective domain (HG) and cognitive domain (CS and HG). For publication, writing the physical manuscript (CS), writing the affective/cognitive manuscript (HG) and preparing for publication (CS).

Study 3: Developed interview guide (CS, HG, LF,ZK, LB and EDM). Piloted focus groups (CS). Conducted 26 focus groups concurrently (CS n=10, HG n=16). Thematic analysis within the child groups aged 7-11 years (CS), aged 5-6.9 years (HG), teachers (CS) and experts (HG). Presented to the wider project team as a means of enabling cooperative triangulation (CS, HG). For publication, writing the manuscript (CS and HG) and preparing for publication (CS).

2.16 Methodological Approaches

The research design of this thesis utilises a mixed methods approach and will contribute to the limited evidence-based research available within the PL field. Study two of this thesis utilises a rigorous quantitative approach by undertaking a systematic review guided by the PRISMA-P guidelines (Moher et al.,2015; Mokkink et al., 2017) reviewing existing assessments within the physical, cognitive and affective domains of PL. In particular, the key aim of the systematic reviews is to expand upon the work previously presented in Edwards (2018a) reporting on existing assessments of PL. Specifically, the work included within this thesis aims to provide much needed detail regarding the psychometric properties, PL alignment and feasibility of existing assessments for use within a primary school setting. As such, the methodological quality of each included assessment tool was appraised using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) risk of bias checklist (Terwee et al., 209; Mokkink et al., 2018). The COSMIN checklist is highly regarded as it was developed in a multidisciplinary, international consensus-study in which 43 experts in health status measurement participated internationally (Mokkink et al., 2012). Furthermore, transparency is a crucial element of conducting systematic reviews (Khan, Kunz, Kleijnen and Antes, 2003) and due to the detailed instruction manual and the background information provided on the development and validation of the checklist this ensured the relevant stages within the review adhered to stringent guidelines, with results recorded in a clear, transparent manner (Terwee et al., 209; Mokkink et al., 2018).

The feasibility, interpretability, cost and ease of use of each instrument was reviewed using a utility matrix based upon work from Klingberg et al (2018). A detailed description of the feasibility concepts, scoring criteria and rationale for rating structure was included within the checklist and this was deemed to be a comprehensive starting point for considering feasibility for assessment tools included within the review. Furthermore, the checklist was

developed with primary schools as a key contextual setting and considered generalist teachers as the assessor which is consistent with setting and age range within this study.

Finally, a bespoke PL checklist was developed to appraise alignment to the concept based on work from Whitehead (2010) and the Australian Spots Commission (ACS) (2017). At the time of development, the ACS were a prominent research group investigating PL and providing informative resources to understand the concept more comprehensively, with a particular focus on the PL domains (Shearer et al., 2018). Further, this work was collated with Whitehead's interpretations of the PL to produce a clear checklist, capturing each PL domain and associated element, thus ensuring transparency in the process of mapping alignment to the concept.

In contrast, study three employed a qualitative technique by using focus groups to collect stakeholder's perceptions of assessment in PL and PE. As perceptions of PL are generally unknown within both teacher and child groups, the selection of focus groups allowed for a much needed detailed and in-depth exploration. Further, through facilitating a conversation, participants can share their experiences and voices in a safe and comfortable environment (Smith and Sparkes, 2016; Gibson 2007; Domville, 2018). To facilitate the focus groups a semi-structured interview guide was developed based on recommendations from Bowen et al., (2009) who suggested several areas of focus when exploring feasibility of new assessments. Three areas were deemed to be of particular importance in line with the aims of the present study; (a) acceptability (to what extent is a new assessment judged as suitable?), (b) demand (to what extent is a new assessment likely to be used?), (c) implementation (to what extent can an assessment be successfully delivered to intended participants?). Recommendations from Bowen et al., (2009) are commonly used within feasibility studies, typically when considering the feasibility for practitioners (Lander et al., 2016; Klingberg et al., 2018). In addition, the framework provided clear direction that

sufficiently assisted with the exploration of research questions within this thesis. Each focus group transcribed was analysed both deductively (using Bowen et al., (2009) as a thematic framework) and inductively, enabling additional themes to be generated (Braun & Clarke, 2006; 2019). Similar to previous studies in related areas (Morley, Van Rossum, Richardson & Foweather, 2018; Ni Chróinín & Cosgrave, 2013), the use of focus groups allowed for deep and meaningful insight into the perspectives of participants, which subsequently allowed for the construction of meaningful themes.

The final study of this thesis is arguably the most impactful as it provides a detailed examination of all qualitative and quantitative data collected throughout the project in order to produce informed recommendations for PL assessment within the 7-11 age range. The inclusion of both qualitative and quantitative research within this has provided the opportunity for contextual information to be collected to expand on the quantitative findings within the systematic reviews. Furthermore, embracing a mixed methods approach throughout has allowed 'a more complete picture' of PL assessment within the 7-11 age range (Denscombe, 2008).

Chapter Three
Global Interpretations
Of PL

The main outcomes of this study have been published in: Shearer, C., Goss, H. R., Edwards, L. C., Keegan, R. J., Knowles, Z. R., Boddy, L. M., ... & Foweather, L. (2018). How is PL defined? A contemporary update. *Journal of Teaching in PE*, *37*(3), 237-245. Due to the university thesis submission regulations this study must be submitted as it is published. Therefore, the strengths and limitations of this study are included within the synthesis chapter (Chapter 7) of this thesis.

3.1 Thesis study map: Study 1

Study	Objectives
Study 1: Global interpretations of PL	• To collate, compare, and critically review existing definitions of PL from leading organisations implementing PL agendas around the world

3.2 Introduction

Over the past 20 years, the invigoration of research regarding physical activity and physical education has generated a greater understanding of both their importance, and how they should be promoted (Allan, Turnnidge, & Côté, 2017). "Physical literacy" has subsequently emerged as a concept that captures both the desire to participate in physical activity, as well as gaining meaningful, fulfilling experiences through doing so. The concept was initially proposed by Whitehead (2001, 2010), in response to concerns as to the direction of physical education and the alarming levels of physical inactivity across the lifecourse (Hallal et al., 2012). Physical literacy has been presented as a "longed for" approach, that values our physical existence (Lundvall, 2015, p. 116). Crucially, it redefines how physical activity is understood, and places importance on the holistic development of an individual's physical potential (Whitehead, 2010). This approach appears to have wide appeal (Jurbala, 2015; Tremblay & Lloyd, 2010), with nations from across the world embracing physical literacy to better promote the health, productivity, and happiness of their citizens. The concept of physical literacy is, however, often interpreted differently between and within these countries (Edwards, Bryant, Keegan, Morgan, & Jones, 2017), leading to concerns that the concept is becoming lost, confusing, or that it is being implemented in ways that are inconsistent with its own core tenets (Jurbala, 2015). As such, researchers have endeavoured to elaborate on what the concept means and how it can be applied in practice. Nevertheless, research published on the concept of physical literacy has provided a diverse array of perspectives (Dudley, Cairney, Wainwright, Kriellaars, & Mitchell, 2017; Edwards et al., 2017), which will be further explored within this paper.

The Origins of Physical Literacy

According to Whitehead (2001), physical literacy is derived from the philosophical concepts of monism, phenomenology and existentialism. "Monism" is the belief that the mind and

body are interdependent and indivisible (Whitehead, 2007). "Existentialism" proposes that every person is an individual as a result of their interactions (Whitehead, 2007). Similarly, "phenomenology" proposes that individuals are formed through their experience of these interactions, and suggests that perception, through our embodied nature, forms unique perspectives in how individuals view the world (Whitehead, 2007). As such, under these assumptions, at the core of physical literacy, individuals will have: (a) a unique interpretation of the physical world, (b) embodiment within this world based on their own experiences and perceptions, and (c) their physical and mental being viewed as an indivisible, mutually enriching whole. It should be noted, however, that each of the philosophical concepts of monism, existentialism, and phenomenology were originally proposed as self-contained approaches to the philosophy-of-science, and not intended for mixing (Grix, 2002).

Whitehead's intention (cf. Whitehead, 2010), by invoking these stances, was to transform physical literacy into an inclusive and holistic concept, focussed on the individual-in-the-world, and her/his experiences. Whitehead (2010) argued that one cannot fully understand or appreciate the true nature of physical literacy without first grasping its philosophical concepts. Yet for many, the detailed and complex philosophical groundings of physical literacy present a barrier to clarity and understanding (Jurbala, 2015). For researchers seeking to explain the concept, there must be some understanding of the philosophical assumptions in order to validate predictions, and this should be articulated. Recent analysis in the related domain of sport and exercise psychology has suggested that the lack-of-willingness to discuss and consider philosophical underpinnings is the cause of many current discrepancies, disagreements, and plateaus in progress (Hassmén, Keegan, & Piggott, 2016).

A definition is, or should aim to be, inextricably linked to its underpinning philosophical assumptions (Dennett, 1995). Whitehead has been proactive in seeking to

refine and improve the definition of physical literacy since she first proposed the concept in 1993 (Whitehead, 1993), often through consensus-seeking exercises within the International Physical Literacy Association (IPLA). For example, in 2010 physical literacy was defined as: "appropriate to each individual's endowment, physical literacy can be described as the motivation, confidence, physical competence, knowledge, and understanding to maintain physical activity throughout the lifecourse" (Whitehead, 2010, p. 11). In 2013, Whitehead had described physical literacy in the International Council for Sport Science and Physical Education (ICSSPE) bulletin as "the motivation, confidence, physical competence, knowledge, and understanding to value and take responsibility for maintaining purposeful physical pursuits/activities throughout the lifecourse" (Whitehead, 2013b, p. 29). Following discussions and refinements, the definition was recently changed on the IPLA website, to read as follows: "the motivation, confidence, physical competence, and knowledge and understanding to value and engage in physical activity for life" (IPLA, 2017). While there have been three iterations of the definition since 2001, Whitehead and her colleagues at the IPLA have always retained the elements of motivation, confidence, physical competence, knowledge, and understanding. Another constant throughout Whitehead's definitions is the notion that the concept is applicable throughout the lifecourse. Nevertheless, the evolving nature of the definition may be a pivotal consideration in illustrating how individuals who approach physical literacy as a new/novel concept may be left bewildered in their search for a definitive definition as arguably, none exists at this time.

Generally, good science is embodied by debate, discussion, and a willingness to evolve and progress ideas (Popper, 1957) and, in this respect, physical literacy is thriving. The following sections will demonstrate that while there may not be a correct or true definition, as both consensus and evidence are currently lacking (Jurbala, 2015), instead there are – or should be – transparent approaches (Edwards et al., 2017). This paper aims to collate,

compare, and critically review existing definitions of physical literacy from leading

organisations implementing physical literacy agendas around the world. This process will

thus facilitate the positioning and contextualisation of various policy frameworks,

measurement and assessment approaches, and intervention data and results. Each will be

discussed with respect to its specific underlying definition and conceptualisation. Common

themes and differences will then be discussed, as well as origins for these differences. While

other papers have sought to critically appraise varying concepts (Robinson & Randall, 2017),

or offer their own interpretations (Chen, 2015), the aim of this paper is to clearly identify,

articulate, and compare the various approaches of each group, united under the label of

physical literacy.

3.3 Methods

Members of the IPLA (n=4) were contacted via email in Spring 2017 and asked to identify

leading organisations/groups working within the physical literacy community. Physical

literacy is a relatively novel concept with almost all organisations/groups using freely

available online platforms to share research and express definitions and interpretations.

Working with these experts allowed access to definitions produced both inside and outside of

the traditional academic publishing distribution channels. In tandem, the references of a

recent systematic review of definitions, foundations, and associations of physical literacy

(Edwards et al., 2017) were also checked to ensure all relevant organisations/groups and

resources were identified. The websites and publicly available material from each

organisation/group were searched to capture information regarding the definitions and

theoretical/conceptual underpinnings of physical literacy being operationalised

internationally.

3.4 Results

We identified that there are seven prominent groups currently working to promote and develop physical literacy, each operating with at least one identifiable definition. The groups included research teams, government organisations (national or state), not-for-profit and corporate groups, or multi-sector partnerships spanning all of these. These organisations/groups use online platforms to share research and present definitions and interpretations of the concept and these were used to gain insight. Definitions and interpretations of physical literacy from each of these seven groups are presented according to country of origin in Table 3.1.

Table 3.1 International Definitions of Physical Literacy

Group	Country of Origin	Reference/ Web link	Adopted Definition of Physical Literacy
International Physical Literacy Association (IPLA)	UK	Whitehead (2017) https://www.physical-literacy.org.uk/	Physical literacy can be described as the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for engagement in physical activities for life
Sport Wales	Wales (UK)	Sport Wales (2017) http://physicalliteracy.sportwales.org.uk/en/	Physical Skills + Confidence + Motivation + Lots of opportunities = Physical Literacy
Physical and Health Education (PHE) Canada	Canada (Montreal)	PHE Canada (2017) http://www.phecanada.ca/programs/physical-literacy/what-physical-literacy	Individuals who are physically literate move with competence and confidence in a wide variety of physical activities in multiple environments that benefit the healthy development of the whole person
Canadian Sport for Life (CS4L)	Canada (Toronto)	CS4L (2017) http://sportforlife.ca/qualitysport/physical-literacy/	Physical literacy is the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for engagement in physical activities for life
Society of Health and	United States of America	Mandigo, Francis, Lodewyk & Lopez	Physical literacy is the ability to move with competence and

Physical Educators (SHAPE)		(2012) http://www.shapeamerica.org/events/physicalliteracy.cfm	confidence in a wide variety of physical activities in multiple environments that benefit the healthy development of the whole person
Sport New Zealand	New Zealand	Sport New Zealand (2015) http://sportnz.org.nz/about-us/who-we-are/what-were-working-towards/physical-literacy-approach	The motivation, confidence, physical competence, knowledge and understanding required by participants that allows them to value and take responsibility for engaging in physical activity and sport for life
Australian Sport Commission	Australia	Australian Sports Commission (2017) http://ausport.gov.au/physical_literacy	Four defining statements: 1.Core / process - Physical literacy is lifelong holistic learning acquired and applied in movement and physical activity contexts 2.Components / constructs - It reflects ongoing changes integrating physical, affective (subsequently renamed 'psychological'), cognitive and social capabilities 3.Importance - It is vital in helping us lead healthy and fulfilling lives through movement and physical activity 4.Aspiration / product - A physically literate person is able to draw on their integrated physical, affective, cognitive, and social capacities to support health promoting and fulfilling movement and physical activity - relative to their situation and context

United Kingdom (UK)

The IPLA is a leading advocacy group for physical literacy in the UK, having been established as a UK charity in 2014, whereupon Margaret Whitehead was appointed as the president. The IPLA was formed with the purpose of providing guidance, clarity, and consistency regarding physical literacy. At the time of this study, the IPLA promoted their definition of physical literacy through their website (www.physical-literacy.org.uk), as well as delivering training programmes to practitioners and hosting an annual conference. Nonetheless, there was a lack of research published by the association, and despite being named the "International Physical Literacy Association," the group is predominantly connected with UK partners and focused on promoting physical literacy within the UK.

Despite the establishment of the IPLA, different definitions and interpretations of physical literacy had been utilised across UK countries (England, Wales, Scotland, and Northern Ireland). The importance of physical literacy for children and young people was first affirmed within national government policy and strategy in England in "Sporting Future: A New Strategy for an Active Nation" (Sport England, 2016). In response, Sport England – a non-departmental public body tasked by Department for Culture Media and Sport with increasing population levels of participation in physical activity in England – had identified "increasing the percentage of children achieving physical literacy" as a key performance indicator within their 2016-2021 strategy (Sports England, 2016, p. 20). The Youth Sport Trust, in partnership with Sport England, Association for Physical Education, Sports Coach UK, and County Sports Partnership Network had created a Primary School Physical Literacy Framework, detailing the role of school physical education (PE), extra-curricular activities, and competitive sports. Within this framework physical literacy was defined as the "motivation, confidence, physical competence, knowledge, and understanding that provides children with the movement foundation for lifelong participation in physical activity" (Youth Sport Trust,

2013, p. 1). Although similar to the previously discussed Whitehead definition, the additional outcome of movement foundation implied a movement focus within the physical literacy framework. Notably, the IPLA are also not listed as collaborating or endorsing this framework.

In Wales, the devolved Welsh Government (Llywodraeth Cymru) prioritised physical literacy at a policy level considerably earlier than England, with physical literacy highlighted as an opportunity to enable lifelong participation in sport and physical recreation. As such, recommendations to raise the status of physical education to become a core subject in Wales - alongside mathematics, English, Welsh, and science - were proposed (Schools and Physical Activity Task and Finish Group, 2013). At the time of publication, the physical literacy definition adopted by Sport Wales displayed similarities to the definition put forward by Whitehead and the IPLA, but instead, it was articulated in the form of an equation: "Physical Skills + Confidence + Motivation + Lots of opportunities = Physical Literacy" (Sport Wales, 2017). In turn, the Sport Wales definition was an attempt to translate the complex physical literacy concept into one that the general public could easily interpret. In line with Whitehead's approach, Sport Wales advocated the notion of physical literacy as a journey throughout life through their interactive website (http://physicalliteracy.sportwales.org.uk/en/) that displayed physical literacy in relation to different life stages. Further, in 2014, approximately £1.78 million ($2.3 million) was invested by the Welsh government into the "Physical Literacy Programme for Schools." The program was a targeted intervention programme that aimed to develop young people along their physical literacy journey. The programme had a political agenda of improving young people's engagement and confidence in secondary schools and reducing the impact of deprivation on academic attainment (Sport Wales, 2017). More recently, upcoming curricular changes in Wales were implicitly aligned with the concept of physical literacy, whereby

physical education will be part of the "health and well-being area of learning and experience" that aims to develop "healthy and confident individuals" (Donaldson, 2015, pp. 45-46).

Canada

As a nation, Canada is often praised for being a strong advocate and leader of physical literacy through its implementation of well-funded programmes and strategies within national sport systems (Allan et al., 2017). There are many groups across Canada's provinces and territories using the term physical literacy, with varying definitions and interpretations of the concept. Two leading government funded groups that work to promote physical literacy on a national scale are Canadian Sport for Life (CS4L) and Public Health and Education Canada (PHE Canada). There are also regional groups dedicated to physical literacy research, such as the Healthy Active Living and Obesity group and the Pacific Institute for Sporting Excellence.

Initially a range of physical literacy definitions were developed in Canada, often adapted from the Whitehead (2010) original definition to suit the needs of specific organisations. The Whitehead (2010) physical literacy definition is – in some capacity – recognised or endorsed by each research team or organisation. Nevertheless, in 2015, discourse within the physical literacy community – surrounding concerns for the divergence in approaches and foci of programme – prompted the creation of a consensus statement within Canada. The purpose of the statement was to provide clarity for the development of policy, practice, and research. The consensus statement was a collaborative process and authors of the statement included: ParticipACTION, Sport for Life Society, the Healthy Active Living and Obesity Research Group at the Children's Hospital of Eastern Ontario Research Institute, Physical and Health Education Canada (PHE Canada), Canadian Parks and Recreation Association, and the Ontario Society of Physical Activity Promoters in Public Health (CS4L, 2015). The IPLA

definition (IPLA, 2017) informed by Whitehead (2013b; the motivation, confidence, physical competence, knowledge and understanding to value and engage in physical activity for life) was endorsed within the consensus statement as the definition of physical literacy (CS4L, 2015, p. 1).

Despite the generation of this consensus statement, the previous definitions from these organisations were often referred to in practice and the primary sources available to interested parties searching the internet (Hyndman & Pill, 2017). The prevalence of these competing approaches leads to the continued confusion and disagreement within the physical literacy community (Robinson & Randall, 2017). For example, in 2009, PHE Canada, a leading professional organisation for physical education teachers, released a physical literacy positioning paper using the following working definition: "Individuals who are physically literate move with competence and confidence in a wide variety of physical activities in multiple environments that benefit the healthy development of the whole person" (Mandigo, Francis, Lodewyk, & Lopez, 2012, p. 6). This definition was displayed on the PHE Canada (2017) website (http://www.phecanada.ca/programs/physical-literacy), however, at the same time the IPLA definition was also endorsed with reference to the consensus statement.

In addition to PHE Canada's approach, The Sport for Life Society (previously Canadian Sport For Life) endorses the IPLA definition of physical literacy, alongside the description: "Physical literacy is the mastering of fundamental movement skills and fundamental sport skills" (The Sport for Life Society, 2017). In 2016, The Sport for Life Society registered "60 Minutes Kids Club," which became "Physical Literacy for Life" (PLFL, 2017). PLFL aimed to advance physical literacy in the health, recreation, and education sectors, with the aspiration "to develop physical literacy in all Canadians" (PLFL, 2017, p. 1). Again, the materials accompanying this site reiterated the IPLA 2014 definition of physical literacy, alongside the full 2015 consensus statement, although it has been debated whether this

acknowledgement was translated in practice (Robinson & Randall, 2017). For example, in 2014, physical literacy was adopted as one of the 10 key factors influencing the CS4L model of Long Term Athlete Development (CS4L, 2015). This model became a popular and influential approach often deployed in relation to physical literacy in Canada (Robinson & Randall, 2017). The model evolved to try to acknowledge the wide variety of factors that influence physical literacy, and in turn athletic development, including a range of skills and environments. As an internationally recognised talent development model, this performance-driven approach to physical literacy received global attention (Allan et al., 2017). Nevertheless, although CS4L adopted the IPLA definition of physical literacy, strategies intended to promote physical literacy within the Long-Term Athlete Development model largely focussed on physical skills and motor development (Allan et al., 2017) and as the popularity of this model grew, so too have criticisms regarding whether the model truly acknowledges the holistic nature of physical literacy (Robinson & Randal, 2017).

United States

At the time of our sampling, physical literacy in America was supported by The Society of Health and Physical Educators (SHAPE America) as a part of the National Standards and Grade Level Outcomes for K-12 Physical Education (Moreno, 2013). In 2013, SHAPE America defined physical literacy as "the ability to move with competence and confidence in a wide variety of physical activities in multiple environments that benefit the healthy development of the whole person" (Mandigo et al., 2012, p. 6; SHAPE America, 2014, p. 4). This definition was the same as that utilised by PHE Canada, and physical literacy is outlined as the goal for both physical and health education, highlighted through the campaign 50 Million Strong which reflected SHAPE America's commitment to put all children on the path to health and physical literacy by 2029 (Jefferies, 2016).

In 2015, The Aspen Institute (an education and policy studies organisation) was commissioned by SHAPE America to produce the document: "Physical literacy in the United States: A model, strategic plan, and call to action" (The Aspen Institute, 2015). Alongside the SHAPE America website, the Aspen Institute developed further resources via their "Physical Literacy: Project Play" website which defined physical literacy as "the ability, confidence, and desire to be physically active for life" (The Aspen Institute, 2013), thus deviating quite significantly from the SHAPE America definition. Crucially, this wording removed the element of knowledge and understanding from Whitehead's definitions, although it could be argued that this was in an attempt to simplify the definition in order to engage youth populations. Both Physical Literacy: Project Play (The Aspen Institute, 2013) and SHAPE America are initiatives for school-aged children, so will undoubtedly focus on children and young people.

SHAPE America asserted that physical education "develops the physically literate individual through deliberate practice of well-designed learning tasks" (SHAPE America, 2017, p. 1). In 2014, the term "physically educated" was replaced with "physically literate" in the National Standards and Grade Level Outcomes for K-12 Physical Education (SHAPE America, 2014). This was critiqued by Lounsbery and McKenzie (2015) and it was reported that this change occurred without the consultation of the physical education profession. It was also argued that there appeared to be little difference between the definitions of physical education and physical literacy. This argument was echoed by Hyndman and Pill (2017), who argued that the substitution and interchangeable use of physical education for physical literacy has led to "definitional blurring."

New Zealand

Sport New Zealand is a government-funded agency that supports and funds local, regional, and national organisations working to promote grassroots and elite sports throughout New Zealand. The 2015-2022 Community Sports Strategy (Sport New Zealand, 2015), which followed the first national strategy published in 2009, highlighted physical literacy as a key focus area for young people within New Zealand. To guide this focus area, Sport New Zealand (2015) published a document titled Physical Literacy Approach - Guidance for Quality Physical Activity and Sport Experiences, wherein they used Whitehead's (2013b) definition of physical literacy: "the motivation, confidence, physical competence, knowledge and understanding required by participants that allows them to value and take responsibility for engaging in physical activity and sport for life" (Sport New Zealand, 2015, p.1). Sport New Zealand reasoned that although they wanted to be a successful sporting nation, they require a participant-focused physical literacy approach to community sport. This approach took a holistic view of the participant, considering their physical, social and emotional, cognitive, and spiritual needs (Sport New Zealand, 2015). The inclusion of a spiritual aspect to their interpretation of physical literacy reflected the important spiritual facets of the Maori culture, which is specific to, and has great importance within New Zealand culture and society. Further, Sport New Zealand outlined their vision, provided information regarding physical literacy, and considered the needs and considerations of various life stages. This document (Sport New Zealand, 2015) gave significance to the "lifecourse," in line with Whitehead's (2010) definition, through a section called "traveling through life" where physical literacy was considered in regard to each life stage (i.e., from early years through to seniors), thus promoting a holistic and inclusive approach to physical literacy. The most recent annual report from Sport New Zealand targets improving physical literacy in children between 2017 and 2020 (Sport New Zealand, 2016).

Australia

The first Australia-wide curriculum for Health and Physical Education (HPE) was released to Australia's states and territories and their respective education systems in 2015. Although the HPE documents did not make an explicit reference to physical literacy, there were strong alignments between particular interpretations of physical literacy and aspects of the HPE curriculum; for example, the aim of the curriculum is to provide the basis for developing knowledge, understanding, and skills for students to lead healthy, safe and active lives (Australian Curriculum, Assessment and Reporting Authority - ACARA, 2016). The concept of physical literacy was specifically mentioned in the document titled Getting Australia Moving, which was commissioned by the local state government in the Australian Capital Territory (Keegan, Keegan Ordway, Daly, & Edwards, 2013). During this time, the University of Canberra's physical literacy research group was arguably the leader of physical literacy within Australia (The Aspen Institute, 2015), aiming to improve the physical literacy of Australian children through school physical education and sport, community linkages, and the development of resources such as web apps and task-cards for teachers.

In May 2016, the Australian Sports Commission recruited a team of researchers to produce, for Australia, a physical literacy definition, standards framework, assessment guidelines, and implementation guidelines. The core researchers in the team conducted a wide-ranging literature review of physical literacy, followed by expert panel meetings, and a Delphi consultation process involving three rounds of Delphi surveys to pursue consensus (Australian Sports Commission, 2017). Following this process, it was agreed that physical literacy should be theoretically separable from physical activity, a so-called double dissociation wherein a person could be high or low in both, separately, or together. The group agreed on a set of defining statements making it clear that each individual has the potential to learn through participation in physical activity and that potential can be developed to a level where it is self-perpetuating. In the end, there were four defining statements issued by the

Australian Sports Commission, with between 94-100% consensus recorded from an expert group of 18 leading researchers. The four defining statements were: (a) Physical literacy is lifelong holistic learning acquired and applied in movement and physical activity contexts (Core/process; 94% consensus); (b) It reflects ongoing changes integrating physical, affective (subsequently renamed psychological), cognitive, and social capabilities (Components/constructs; 94% consensus); (c) It is vital in helping us lead healthy and fulfilling lives through movement and physical activity (Importance; 100% consensus); and (d) A physically literate person is able to draw on their integrated physical, psychological, cognitive, and social capacities to support health promoting and fulfilling movement and physical activity – relative to their situation and context – throughout the lifespan (Aspiration/product; 94% consensus).

Central to these defining statements was the clarification that whole-person, holistic development spans four key learning domains: the physical, affective, cognitive, and social (Australian Sports Commission, 2017). The physical domain included physical competence, motor skills, health- and skill-related fitness, technique and psychomotor skills. The affective (subsequently 'psychological') domain concerned itself with one's experiences of internal signals such as fatigue and exertion, as well as motivation, confidence, self-esteem and engagement. The cognitive domain covered conscious and unconscious knowledge and understanding, including problem-solving and decision-making, awareness of rules and tactics, appreciation of healthy and active lifestyles, and processing of feedback and reflection. The social domain included leadership, understanding ethical principles, working with peers, coaches, teachers and more, treating others with sensitivity and effective communication. The group emphasised that development and learning must be "integrated across" all four domains, and not merely focussing on the physical. It is early days for this new approach, using defining statements rather than a singular definition, but the work has

been well received in stakeholder focus groups and has support from the Federal government, including ongoing funding of the Australian Sports Commission's work in this area across Australia.

3.5 Discussion

The current paper has endeavoured to collate, compare, and critically review the current understandings of physical literacy internationally. We have identified seven established and prominent groups, and have provided an overview of those groups operating with the term physical literacy. The following discussion will critically review these by identifying common themes and issues regarding the definitions used by these groups, exploring potential reasons for these issues, and pointing out the implications this has for the future of physical literacy.

Global Differences

In articulating her views on the concept of physical literacy, Whitehead (2010) was clear that there are good reasons to expect different approaches to physical literacy. The underlying philosophy (or philosophies) she argued as being central considerations denoted that the unique personal experience, unique personal capabilities at any point in time, and unique social and environmental contexts all necessitate a context-specific approach. International differences in the interpretation and operationalization of physical literacy are expected, indeed needed, in order to create meaning and cultural relevance. The influence of culture was extensively discussed by Whitehead (2010) who identified that "specific expression (of physical literacy)… will be particular to the culture in which they live" (p. 12). Although physical literacy is proposed to be a universal and inclusive concept, there is a debate as to how much tailoring the socio-cultural context should necessitate, and this is referred to throughout Whitehead's book (2010). Initially, it was assumed that the differences in interpretation could stimulate the implementation of physical literacy in practice and allow it to flourish within a variety of settings, ultimately, leading not only to different approaches to applied practice, but also different definitions of physical literacy. As a consequence, however, some have argued that this diversity in definitions has generated a level of

inconsistency and conflict within the physical literacy community (Dudley et al., 2017; Jurbala, 2015; Tremblay & Lloyd, 2010).

Each of the seven organisations, discussed above, have adopted their own definition(s) of physical literacy. With the exception of SHAPE America, these groups are non-governmental public sports bodies. While the growing interest from international organisations aiming to promote physical literacy is promising, it should be noted that these organisations each have their own specific purposes, philosophies, expertise, and funding priorities in order to promote the concept within their communities. These contextual constraints then influence associated characteristics, descriptors, objectives, methodologies, programmes, and evaluations of physical literacy, perhaps perpetuating the issues that form the focus of the current paper.

The Canadian consensus statement (CS4L, 2015) aimed to decide upon a single definition as, even within one country, the interpretations of physical literacy were notably different across provinces. The Canadian consensus statement went some way towards unifying a physical literacy approach, yet there is a marked difference between endorsing a definition and appropriately operationalising said definition (Edwards et al., 2017). It is unclear, however, what meaningful difference this consensus achieved in terms of changes to practice and approaches, with conflicting definitions presented alongside the 'agreed' one. More substantive, transparent, and scientific processes may be required in order to develop and agree on a robust working consensus regarding the definition and meaning of physical literacy.

Philosophy within the definition. The philosophy underpinning the physical literacy concept and its holistic nature is arguably what makes the concept unique. Whitehead has consistently noted that philosophy is the vital foundation behind physical literacy and one cannot truly

understand physical literacy without embracing its philosophical roots (2001, 2007, 2010, & 2013b). Nevertheless, the philosophy surrounding physical literacy programmes was often ill-aligned or simply missing, both in research and practice (Edwards et al., 2017). For example, SHAPE America (2017) and Sport Wales (2014) may have neglected the lifelong experience in their materials, as their focus at the time was on school-aged populations. Likewise, having historical associations with talent development pathways, The Sport for Life Society (2017) and Sport New Zealand (2016) may have placed higher importance on movement skills rather than valuing the diverse and holistic construction of physical literacy. Yet despite the emphasis on philosophy, Whitehead has never successfully included an acknowledgement of philosophy within the definitions she has developed, or helped to stimulate. This may be a potential reason for the confusion and misinterpretations surrounding the concept.

Defining the Core Elements

While making the concept culturally relevant, some organisations may have deviated from the original Whitehead (2001) definition, which included the four elements of confidence, physical competence, motivation, and knowledge and understanding. For example, CS4L (2015) and PHE Canada (2017) expressed the physical literacy elements as "fundamental movement and sport skills" (CS4L, 2015, p. 1) and "competence and confidence" (PHE Canada, 2017, p. 1). In each case, some of the physical literacy core elements described in Whitehead's definition are omitted; therefore, is the term physical literacy appropriate? Whitehead's definition has taken different forms over the 10 years preceding this analysis, however, it remained consistent in the sense that all four elements (motivation, confidence, physical competence, and knowledge and understanding) were included. Sport Wales (2017) replaced the element "physical competence" from the Whitehead definition with "physical skill." This was seemingly an attempt to translate the core elements into language that can be

easily understood by the general population, thus making it possible to implement within local and education sectors.

Sport Wales (2017, p. 1) added an additional core element, "a range of opportunities" referring to facilities available and the environment facilitating physical activity. By adding this element into the definition, Sport Wales emphasised that physical literacy was not only the responsibility of the individual, but also of parents, teachers, council members, and the community as a whole. Similarly, CS4L (2015), PHE Canada (2017), and SHAPE America (2014) also added this element referring to it as "multiple environments." This aspect was discussed extensively by Whitehead (2001), who sought to clarify what constituted a physically challenging environment, and how a physically literate individual would read the environment. In contrast, however, interacting with the environment was not featured in Whitehead's subsequent definitions (2001, 2007, 2010, 2013a, & 2013b; IPLA, 2017). Recent research by Dudley et al. (2017) identified movement contexts as a significant consideration for policy makers, so much so as to suggest the Whiteheadian definition could beneficially be adapted further to incorporate this crucial element. Interestingly, and in contrast to other groups, Australia's new approach does not mention the four elements of motivation, confidence, competence, and knowledge and understanding. Instead, it has included the components/constructs of physical, affective (subsequently psychological), cognitive, and social capacities (Australian Sports Commission, 2017). The research group reached a consensus that it would be more inclusive and engaging to specify the broader domains as there were concerns that concepts such as motivation and confidence held different meanings to different cultures, between researchers, and versus the wider stakeholder group. This presents an alternative interpretation in approaching physical literacy, which warrants consideration.

A Lifelong Journey

Whitehead (2001, 2010) consistently argued that physical literacy represents a lifelong journey. A recent systematic review of the definitions of physical literacy conducted by Edwards et al. (2017) found "throughout the lifespan" as a core category in defining physical literacy. Within existing literature, they reported the existence of three categories: throughout the lifespan, unique journey, and the Long-Term Athlete Development model. Nonetheless, the systematic review also highlighted physical education as a core category, alluding to the focus that has been placed upon school-aged populations.

Despite most of the groups reviewed advocating Whitehead's definition (2001, 2007, 2010, 2013a, & 2013b; IPLA 2017) to some degree, many groups that have operationalised physical literacy in practice have predominantly focused on school-aged children and young people. This is not surprising, especially as PHE Canada and SHAPE America are organisations formed within the physical education sector. Many of these organisations have received funding from governments who wish to invest in children's health. Particularly within policy, where cost versus benefit must be evidenced, the lack of research to support physical literacy across the lifecourse presents a major barrier. At the time of writing, much of the published literature relating to physical literacy concerned school-aged populations. Within the 2013 special issue on physical literacy published in the Journal of Sport Science and Physical Education, authors admitted many of the articles were school focused (Weinburg, 2013). Likewise, within the current special issue, articles also focus on physical education, as is the mission of the Journal of Teaching in Physical Education. Therefore, in order to generate evidence throughout the lifecourse, relevant and appropriate research from the established contexts of physical education and physical activity should be considered. Nevertheless, physical literacy has only been adopted by policymakers in recent years, and the youth population has evidently been the easiest to access and impact. Perhaps it is too

early to comment on the focus of applied practice. We would suggest that a more holistic approach needs to be taken to consider physical literacy across the lifecourse.

Process Versus Product

An apparent difference when comparing global organisations became the choice of some groups to define a physically literate person as opposed to defining physical literacy. For example, achieving physical literacy in children is a key performance indicator in Sport England's (2016) strategy for physical activity in the UK. Similarly, PHE Canada (2017) described a person who is physically literate in their definition, while SHAPE America identified that physical education is the means "to create the conditions for all youth in the United States to be physically literate by the middle school years" (The Aspen Institute, 2015, p. 11). This process (journey) versus product (outcome/goal) debate became apparent in the work of Keegan et al. (in review), and has led to a core point of difference in the work produced from Australia. The Australian (2017) defining statements differentiate between physical literacy as a process (Statement 1 – Core/process) versus physical literacy as the product/outcome (Statement 4 – Aspiration/product). Different approaches to physical literacy have emphasised an inherent, ongoing potential to learn and develop through movement (process), which has been contrasted against some kind of current physical literacy status (product), which is presented as a desirable level of being physically literate. Concerns remain, however, that discussing physical literacy as an end state, also implies that someone may be physically illiterate, which has been a particular source of contention; Whitehead (2013a) argued that physical illiteracy cannot occur in a living being as human movement potential is necessary for life. Nonetheless, in the book Physical Literacy: Throughout the Lifecourse, Whitehead refers openly to "physically illiterate individuals" (2010, p. 7). In a recent personal communication, Whitehead has expressed frustration at the process versus outcome (versus both) debate. Whitehead has attempted to clarify her view

that although a journey is a process in the interests of seeking a goal, progress on a physical literacy journey depends on the accumulated processes in which the individual is involved (Whitehead, personal comunication, August 14, 2017). Separately, the ongoing process versus outcome (versus both) debate is another core source of disagreement and inconsistencies in definitions, viewpoints, and approaches. Robust and contemporary research on this topic should be published in publically accessible peer-reviewed journals, to engage and render transparent the current debate, thus also stimulating the development of understanding of physical literacy.

Future Implications

This review of the current approaches to defining physical literacy, while not exhaustive, has identified several distinguishable approaches, between and within different countries. For example, in conducting this review we have been made aware of physical literacy programs being conducted in Singapore, Scotland, China, and India. At the time of writing, these programs were not sufficiently developed, or distinguishable from other programs, to warrant a separate analysis. Nonetheless, a common issue experienced by both established and emerging groups working around physical literacy is a lack of empirical evidence (Giblin, Collins, & Button, 2014; Jurbala, 2015). This paucity-of-evidence was a limiting factor in the current paper, as we were only able to include established organisations, all of which existed in English speaking, developed countries. Yet even in these groups, many had an online presence without a peer-reviewed, published evidence-base. Conducting peer-reviewed research and robustly evaluating programmes throughout policy and practice should therefore be a key focus for organisations moving forward.

Crucially, however, when presenting this empirical evidence, understandings of, and assumptions regarding, physical literacy should be clearly presented in order to provide a

frame for interpretations of findings. While the concept and topic of physical literacy appears to hold strong potential – particularly the notion of re-emphasising the holistic, integrated nature of personal development through movement experiences – researchers within the area have increasingly recommended that academics need to focus on clearly articulating aligned definitions, philosophical assumptions, and conceptual frameworks (Dudley et al., 2017; Edwards et al., 2017). Furthermore, with this research transparency, there is also a need for tolerance for differing approaches of physical literacy in order to permit collaborations, sharing, and critical discussions while operationalising the concept (Edwards et al., 2017). This paper demonstrates that different approaches have been adopted towards physical literacy by different groups. Some advocates, often from a specific group promoting a specific approach, are troubled by this divergence in meanings, calling for alignment to agreed core elements of definition wordings. While this paper recognises that there will be different interpretations of physical literacy, it also urges all authors and researchers to clearly articulate their definition, assumptions, and core values when they deliver and report their findings in relation to physical activity and physical literacy.

3.6 Conclusion

A number of international groups, and numerous papers, chapters, and books, have focussed on physical literacy in the recent years. Such is the perceived benefit of physical literacy that within the UK, Canada, USA, New Zealand, and Australia, the term physical literacy has been recently cited within recent national policies. Nonetheless, in order for physical literacy to develop, robust evidence-based research is needed. Within such research, a level of clarity, transparency is needed; and through such clarity and clear evidence, consensus may be pursued regarding the "what and for what" questions (Edwards et al., 2017). To be clear, we do not advocate that each group adopts the same definition a priori, but it must be possible to compare different interpretations and evaluate the effectiveness of measurement/assessment attempts, intervention programmes, and policies internationally. Opportunities for cooperation in promoting physical literacy should continue to be developed, as open discussions could help determine the importance of physical literacy in research and practice (Corbin, 2016). As such, all stakeholders, throughout both academia and applied practices, should seek to clearly and coherently articulate their approach to physical literacy in order to make meaningful differences that stand a chance of significantly advancing the field.

Chapter 4
A Systematic Review of Assessments related to Physical Literacy

4.1 Thesis Study Map

Study	Objectives
Study 1: Global interpretations of PL	• To collate, compare, and critically review existing definitions of PL from leading organisations implementing PL agendas around the world Key findings: • Seven international groups were identified as currently working to promote and develop PL, each operating with at least one identifiable definition of the concept • There is a lack of consensus and clarity between definitions and interpretation of PL, as such transparency is needed when reporting alignment to the varying PL concepts • Whitehead's definition (2001) of PL has been utilised or expanded upon within the differing interpretations, providing the rationale for anchoring the work included within this thesis to Whiteheads 'conceptualisation of PL
Study 2: A systematic review of assessments related to the physical domain of PL among children aged 7-11 years	• To systematically review the academic literature for tools to assess the physical, affective and cognitive domains of PL within children aged 7-11 years. • To explore and critically discuss each assessment tool within the physical, affective and cognitive domains to appraise its (a) psychometric properties; (b) feasibility for use within a primary school setting and (c) alignment to the PL concept.

4.2 Introduction

Previous chapters included within this thesis highlight the overlap between Whitehead's articulation of PL and established research fields, i.e. motor competence, psychology (Gallahue, Ozmun, & Goodway, 2012; Ulrich, 2000; Lubans et al., 2010; Hulteen et al., 2018). As such, it is essential to investigate existing assessments to consider how available tools can be used to support the assessment of PL. In addition, while there are few assessments of PL available the methodological rigour, feasibility and alignment to the concept of each assessment is generally unknown. Therefore, this chapter strives to provide this much needed detail surrounding existing assessments to support the assessment of PL within primary school aged children.

Within the current literature, there is a considerable lack of evidence relating directly to the assessment of PL in children (Cairney at al., 2019; Edwards et al., 2018). Attempting to assess a complex and interchangeable concept such as PL is a challenging process with many traditional assessments such as questionnaires, fitness and functional/fundamental movement tests not deemed appropriate to capture the approach as a whole (Durden-Myers, Whitehead and Pot, 2018; Shearer et al., 2018; Durden-Myers, Green and Whitehead, 2018; Edwards et al., 2017). This is of particular importance as Whitehead (2019) advocates that assessing PL as a whole entity is imperative to aligning with the underpinning philosophy and as such, stresses the importance of equal weighting within the domains.

A recent review from Edwards *et al.* identified 52 assessments of PL and related constructs evaluating these in relation to age group, environment, and philosophy (Edwards et al., 2018). Within the fifty-two assessments that were identified; twenty-two were categorised into the physical domain, with five providing measures for the affective domain and one considering the cognitive domain. Additionally, only one assessment represented all three

domains within the tool. Notably, 66% of assessments were used in children under 12, suggesting this age group as being a key area in existing research. However, within the review, Edwards and colleagues may have missed some relevant assessments as their search terms were limited to 'PL' and did not include search terms specific to each element of PL (i.e., motivation, confidence, physical competence, knowledge and understanding). In addition, while alignment to the philosophy was appraised within the review, the precise methodological process of doing so was unclear. Subsequently, this leaves uncertainty regarding the alignment of each tool with Whiteheads' articulation of the concept and prevents examples of best practice being highlighted. Moreover, the review did not examine the psychometric properties of identified assessments despite validity and reliability of PL assessments being criticised within the literature (Longmuir and Tremblay, 2016). It was also outside the scope of the review to examine the feasibility of using each assessment in practice. This is of particular important as within the primary school context, numerous barriers have been cited regarding effectiveness and feasibility of assessment in PE (Chróinín and Cosgrave, 2013).

Aims and Research Questions

The aim of this study is to systematically review the academic literature for tools to assess the physical, cognitive and affective domains of PL within children aged 7-11 years. Within the literature, there is currently a plethora of assessment information and measurement tools available related to the PL domains, and despite not being marketed as 'PL' tools it would be counterproductive to ignore prior developments in the field of PA assessment (Cairney et al., 2019; Dudley et al., 2017; Barnett et al., 2016; Hulteen et al., 2018; Rudd, Butson and Barnett, 2016). The current study therefore aims to expand on previous work by conducting a wide reaching search of the literature using search terms specific to each PL domain:- **physical**: object control, balance, locomotor, movement skills

(land), movement skills (water), moving using equipment, cardiovascular endurance, muscular endurance, coordination, flexibility, agility, strength, reaction time, speed, power, rhythmic, aesthetic/ expressive, sequencing, movement specific to an environment and progression; **affective**: confidence, motivation, emotional regulation, enjoyment/happiness, empathy, persistence/resilience/ commitment, adaptability, willingness to try new activities, autonomy, comfortable and connected with the world, self-perception/ self-esteem and perceived competence; and **cognitive**: benefits of PA, importance of PA, effects of PA on body, PA opportunities, sedentary behaviour, ability to identify and describe and movement, creativity/imagination, decision making, ability to reflect, tactics, rules, strategy, safety considerations and risk). This study will explore and critically discuss each assessment tool within the physical, affective and cognitive domains to appraise its (a) psychometric properties; (b) feasibility for use within a primary school setting and (c) alignment to the PL concept.

4.3 Methods

The methodology of this study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols (PRISMA-P) (Moher et al., 2015). The protocol information for this review was registered with PROSPERO reference: **CRD42017062217**

Inclusion Criteria

Target groups were comprised of:

1. Typically developing children with a reported mean age or age range between 3-11 years (including overweight and obese children and children from deprived areas)

Studies were included if they:

1. Reported on a field-based assessment tool (qualitative or quantitative), used in the context of PA, sport, PE, active play, exercise or recreation; with an outcome relating to PL (for details, see appendix)

2. Cross-sectional, longitudinal or experimental study design

3. Reported a measurement method relevant to an element of PL

4. Reported an aspect of psychometric testing or theoretical development

5. Published in English and in a peer-reviewed journal

Exclusion criteria

Studies identified through the literature search were excluded if:

1. Assessment tool is not used in the 7-11-year child population

2. Included special populations (i.e. children with DCD, diagnosed with learning difficulty)

3. Lab-based assessment (i.e. VO$_2$ max test, Wingate test)

4. Book chapters, case studies, student dissertations, conference abstracts, review articles, meta-analyses, editorials, protocol papers and systematic reviews

5. Not published in English and not in a peer reviewed journal

6. Aspect of psychometric testing not reported

7. Full text articles were not available

Information sources, search strategy and study selection

Study selection

Relevant studies were identified by means of electronic searches on EBSCOhost, scanning reference lists of included articles, searching author bibliographies and contacting relevant experts (defined as individuals who have experience, knowledge and published work in peer

reviewed journals in each PL element sector). The EBSCOhost platform supplied access to: MEDLINE, PsycINFO, Scopus, Education Research Complete, and SPORTDiscus databases. Each of the databases was searched independently. Publication date restrictions were not applied in any search with the final search conducted on 10th January 2019.

Search strategies used in the databases included combinations of key search terms which were divided into four sections: tool (Assessment OR Measurement OR Test OR Tool OR Instrument OR Battery OR Method OR Psychometric OR Observation OR Indicator OR Evaluate OR Valid Or Reliable) AND context ("PA" OR "PL" OR Play OR Sport OR "PE" OR Exercise OR Recreation) AND population (Child OR Youth OR Adolescent OR Paediatric OR Schoolchild OR Boy OR Girl OR Preschool OR Juvenile OR Teenager) AND PL elements (Motivation OR Enjoyment OR Confidence OR Self Or "Perceived Competence" OR Affective OR Social OR Emotion* OR Attitude* OR Belief* OR Physical* OR Fitness OR Motor OR Movement* OR Skills* OR Technique* OR Mastery OR Ability* OR Coordination OR Performance OR "Perceptual Motor" OR Knowledge OR Understanding OR Value OR Cognition* OR Health OR Wellbeing*). Boolean searches were also carried out using ''AND'' to combine concepts. Following the initial search, the two lead authors (CS and HG) removed all duplicates and screened the titles and abstracts. Only articles published or accepted for publication in peer reviewed journals were considered. A third author (LF) checked decisions, and any disagreements were resolved by discussion and collaboration with all authors. Full-text articles were further evaluated separately for relevance by two authors (CS and HG) and labelled 'yes', 'no', or 'maybe'. The reviewers conferred and, following discussion on any inconsistencies, agreement was reached on all articles.

Data Collection Processes

Two authors (CS and HG) extracted study data relating to: study information (authors, publication date, country and study design), sample description, purpose of study, PL element being assessed (motivation, confidence, physical competence, knowledge and understanding), measurement technique (i.e. interviews, questionnaires, practical trial), outcome variables, quantitative psychometric testing results (COSMIN risk of bias checklist and utility information (Mokkink et al., 2017). Study authors were contacted, where possible, for missing or incomplete data. Both reviewers performed the data extraction process independently, and any discrepancies were resolved through discussion with a third reviewer (LF).

Quality Appraisal

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist was used to evaluate the methodological rigour of assessments (Mokkink et al., 2017). The COSMIN checklist was designed and validated for use in evaluating the rigour of psychometric studies of healthcare instruments. The checklist is of a modular design, which enabled flexibility to suit the needs to the current systematic review. The PROM development, content validity, structural validity, cross-cultural validity, criterion validity, reliability, internal consistency, measurement error, responsiveness and hypothesis testing were appraised with the newly developed COSMIN risk of bias checklist (Mokkink et al., 2017) and subsequently given a rating of; 'very good', 'good', 'doubtful', or 'inadequate' or in either case if not reported 'NR'. The 4-point rating scale and worst score counts method are used throughout. For all measurement properties, two reviewers assessed (CS and HG) the quality separately and determined the consensus ratings in a face-to-face meeting. Additionally, the methodological quality of content validity and PROM development was assessed using the newly developed COSMIN risk of bias checklist for PROMs (Mokkink et al., 2017). The COSMIN guidelines were recently updated (Moher et al., 2015) during the

review process and new guidance regarding the importance of each psychometric property was detailed. According to the updated guidelines, if the original study, associated paper or tool manual does not adequately describe the PROM development process and or aspects of content validity then the tool should not be appraised further. However, to utilise the research already conducted within this review process, this review reports on all 10 psychometric properties included within the original guidelines (PROM development, content validity, internal consistency, reliability, measurement error, structural validity, hypothesis testing, cross-cultural validity, criterion validity and responsiveness).

Table 4.1 Detailed description of rating of measurement properties

Psychometric properties	Definition	Rating	Quality criteria
Content validity	The extent to which the domain of interest is comprehensively sampled by the items in the measurement instrument	+	The target population considers all items in the measurement instrument to be relevant AND considers the questionnaire to be complete
		-	The target population considers all items in the measurement instrument to be irrelevant OR considers the questionnaire to be incomplete
		?	No target population involvement [1]
Structural validity	The degree to which the scores of a measurement instrument are an adequate reflection of the dimensionality of the construct to be measured	+	Factors should explain at least 50% of the variance
		-	Factors explain <50% of the variance
		?	Explained variance not mentioned [1]
Internal consistency	The degree of the interrelatedness among the items	+	(Sub)scale unidimensional AND Cronbach alpha >0.70

		-	(Sub)scale not unidimensional OR Cronbach alpha <0.70
		?	Dimensionality not known OR Cronbach alpha not determined [1]
Cross-cultural validity/ measurement invariance	The degree to which the performance of the items on a translated or culturally adapted measurement instrument is an adequate reflection of the performance of the items of the original version of the measurement instrument	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's R^2 <0.02)
		-	Important differences between group factors OR DIF was found
		?	No multiple group factor analysis OR DIF analysis performed [2,4]
Reliability	The proportion of the total variance in the measurements which is due to 'true' differences between participants	+	ICC OR weighted Kappa r >0.70
		-	ICC OR weighted Kappa r <0.70
		?	ICC OR weighted kappa not reported [1,2,3]
Measurement error	The systematic and random error of a participants score that is not attributed to true changes in the construct to be measured	+	Area under ROC curve is >0.5
		-	Area under ROC curve is <0.5

		?	Area under ROC curve not determined
Criterion validity	The degree to which the scores of an assessment is an adequate reflection of a 'gold standard'	+	Convincing arguments that gold standard is ''gold'' OR alternative measure has been previously validated AND correlation with gold standard OR alternative measure >0.70
		-	Correlation with gold standard OR alternative measure <0.70 despite adequate design and method
		?	No convincing arguments that gold standard is ''gold'' OR alternative measure has been validated OR doubtful design or method [3]
Hypothesis testing for construct validity	The extent to which scores on a particular measurement instrument relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the concepts that are being measured	+	At least 75% of the result is in accordance with the hypothesis
		-	<75% of the result is not in accordance with the hypothesis
		?	No hypothesis defined (by the review team) [1, 2]
Responsiveness	The ability of a measurement instrument to detect important changes over time	+	SDC OR SDC < MIC OR MIC outside the LOA OR RR > 1.96 OR AUC> 0.70
		-	SDC OR SDC> MIC OR MIC equals or inside LOA OR RR <1.96 OR

?   AUC <0.70, despite adequate design and methods

Doubtful design or method [3]

---

*(+ = **positive rating**; - = **negative rating**; ? = **indeterminate rating**) Intraclass correlation (ICC) ; Receiver Operating Characteristic (ROC); Differential Item Functioning (DIF); Smallest Detectable Change (SDC); Minimum Important Change (MIC); Limits of Agreement (LOA); Relative Risk (RR); Area Under the Curve (AUC).*

In addition, Table 4.2 displays a utility matrix developed from combining feasibility elements from two systematic reviews and was used to appraise the feasibility of each assessment tool, including the cost efficiency and acceptability of assessments (Klingberg et al., 2019; Beattie, Murphy, Atherton and Lauder, 2015). Finally, a novel PL checklist, displayed in Table 4.3 was developed by the research team to highlight in each assessment the areas of PL. The checklist was developed after an extensive overview of the international PL literature was conducted (Shearer et al., 2018). Each of the included studies were independently scored by two reviewers (CS and HG) using a standardised process to obtain consistent data across all studies. Conflicts (n=14) were resolved through discussion with the review team (CS, HG, LF) until consensus was reached.

Table 4.2 Detailed description of rating of feasibility concepts

Rating		Excellent (****)	Good (***)	Fair (**)	Poor (*)
Cost Efficiency	How long does an assessment take to complete?	<15 min	<30 min	30-60 min	>60 min
	How much space is needed to administer an assessment?	Less than 6 metres, a corner of a room	6-10 metres a standard room	10-20m (primary school sports hall)	20m+ (Secondary school sports hall requirement)
	What equipment is required to administer an assessment?	Equipment likely to be present in a typical school	Some extra equipment or resource required would be additional that what is typically present (primary school)	Most of the equipment required would be additional that what is typically present (primary school)	All equipment required to would be additional that what is typically present (primary school)
	What qualification is required to administer an assessment?	Able to be administered by any school staff	Able to be administered by qualified teacher	Able to be administered by PE/Sport specialist	Requires researcher with specific higher qualifications
	What training is required to administer an assessment?	Little or no additional training required	Some additional training required (less than half a day)	Further additional training required (half a day to one and a half days)	Significant training required (more than one and a half days)
Acceptability	Is there evidence of participant understanding?	Investigation of participant understanding (evidence from	Estimated evidence of participant understanding (evidence from source	Participant understanding not explicitly stated but can be assumed	No evidence of subject understanding

	participants)	other than participant)		
How many assessments are not completed?	Low number of missing items (<10%) and adequate response rate (>40%)	High number of missing items (>10%) and an adequate response rate (>40%)	Low number of missing items or poor (<10%) and an adequate response rate (<40%)	High number of missing items (>10%) and poor response rate (<40%)

Table 4.3 Physical literacy 'sub-elements' identified from literature collated in Study One (Chapter Three)

Affective domain	Physical domain	Cognitive domain
Confidence	Object control	Benefits of physical activity
Motivation	Balance	Importance of physical Activity
Emotional regulation	Locomotor skills	Effects of physical activity on body
Enjoyment/happiness	Movement skills- land	Opportunities to be active
Empathy	Movement skills-water	Sedentary behaviour
Persistence/resilience/ commitment	Moving using equipment	Ability to identify and describe and movement
Adaptability	Cardiovascular endurance	Creativity and imagination in application of movement
Willingness to try new activities	Muscular endurance	Decision-making (ability to think, understand and make decisions, knowing how and when to perform)
Autonomy	Coordination	Appropriate movement strategies that a situation or environment requires
Comfortable and connected with the world	Flexibility	Ability to reflect and improve own performance, including setting optimal challenges
Self-perception// self-esteem	Agility	Tactics, rules and strategy
Perceived physical competence	Strength	Action planning and outcome expectations
	Reaction Time	Safety considerations and risk
	Speed	

Power

Rhythmic ability

Aesthetic/ expressive
ability
Sequencing

Specific to an
environment
Progression

Adapted from Whitehead, 2010; Whitehead, 2013; Dudley, 2015; Longmuir et al., 2015; Longmuir & Tremblay, 2016; Edwards et al., 2017; Keegan et al., 2019)

4.4 Results

An overview of the search process is provided in Figure 4.1. The search strategy resulted in a total of 7530 initial results. Reference checking resulted in three additional eligible studies. After screening of titles and abstracts, 381 articles were retrieved for full text reading. After full text screening was completed a total of 130 studies were included in this review. In total, 65 studies were relevant to the physical domain, 58 to the affective domain and 7 to the cognitive domain. Within the 7-11 age range, 6 tools assessed elements within the cognitive domain, 33 tools assessed elements within the affective domain and 21 tools assessed elements within the physical domain.

4.5 Results of the Physical Domain

The physical domain results include 65 studies providing information for 21 assessment tools within the 7-11 age range. Table 1.5 includes the study characterises and details information relating to geographical location, setting, age range and scoring. The studies were conducted within the USA (n=5), Canada (n=4), Australia (n=1) and Europe (n= 11). Notably, two of the Canadian assessment tools are marketed as PL assessments, specifically the Canadian Assessment of PL (CAPL-2) (Longmuir et al., 2018) and PL Assessment in Youth (PLAYfun) (Cairney et al., 2018). Assessments were typically administered within the gym hall or an onsite sports facility within the school setting (n=20). However, one assessment - The PA Research and Assessment tool for Garden Observation -utilised a different setting and was administered in a community garden (Myers and Wells, 2015). The age range for each tool varied across the review. Notably, some tools (n=2) reported a crossover between age ranges and had specific tasks/instructions for both key stage one (5-7 years) and key stage 2 (7-11years) children (Cepero et al., 2013; Fransen et al., 2014). Additionally, each tool utilised a form of numerical scoring, such as time taken to complete the assessment, awarding levels and distance covered

Figure 4.1 PRISMA flow diagram showing the process of identification and selection (Moher et al., 2009)

Identification

7530 references imported for screening

3 additional resources identified through other sources

7177 after duplicates (n=356) removed

Screening

7117 records screened

6796 excluded on title and abstract

381 full-text studies screened

Eligibility

130 studies included

Cognitive n=7

Affective n=58

Physical n=65

251 full-text studies excluded

Incorrect assessment outcome n=75

Not used with children with a mean age between 7-11 n=67

No assessment of psychometric properties or systematic theoretical development n=52

Book chapter, case study, conference abstract, dissertation, review articles, meta-analysis, protocol, or editorial n=20

Incorrect context n=16

Full text not available n=17

Duplicate publication n=3

Clinical population n=1

Not reporting an assessment n=3

Not English language n=1

Included

Assessments related to the 7-11 age range

Cognitive n=6

Affective n=33

Physical n=21

Table 4.3 The study characteristic information for each assessment within the physical domain

Assessment and country of origin	Participant n, gender (%) (age range; mean age)	Purpose/ use of assessment	Scale Design	Scale Scoring
ALPHA *Spain* *(España-Romero et al., 2010)*	N=58, NR (6-11; NR)	Fitness assessment	Pubertal status Weight and Height Waist circumference Skinfold thickness (triceps and subscapular) Hand grip strength Standing long jump 4x10m shuttle run test 20m shuttle run test	Individual scores for each test: if the student would not perform the task by selecting a reason: 1=shyness, 2=lack of motivation
Athletic Skills Track (AST) ½ *Netherlands* *(Hoeboer, De Vries, Krijger-Hombergen, Wormhoudt, Drent, Krabben & Savelsbergh, 2016)*	N=463, NR (6-12; NR)	FMS	The tracks consisted of a series of fundamental motor tasks (n = 10)	Time taken to complete each track
Bruininks–Oseretsky Test of Motor Proficiency (BOTMP-SF) *Canada* *(Fransen et al., 2016)*	N=590, 45.4% girls (9-11; 11.46± 1.46)	Motor competence	Consists of 4 motor area composites; fine manual control, manual coordination, body coordination, strength and agility	Total point scores; standard scores or percentile ranks

CAPL-2 *Canada* (Tremblay et al., 2018)	N=10034, 50.1% girls (8-10;10.1±1.17)	PL	CAMSA PACER (10m or 25m) Isometric plank hold	CAMSA: The time required to complete the course was recorded, and then converted to a point score (range 1–14). The quality of each skill was scored as either performed (score of 1) or not observed (score of 0)
EUROFIT *Norway* (Fjørtoft, 2000)	N=75, 49.3% girls (5-7; 6.1± NR)	Fitness assessment	6 minute run test Arm pull(or hand grip) Standing broad jump (or vertical jump) Bent arm hang Sit-ups Sit and reach Plate tapping Shuttle run (10x5 meters) (or 50m sprint) Flamingo balance	Highest score for each assessment recorded
FG-COMPASS *USA* (Furtado & Gallagher, 2012; Calatayud, Martín, Colado, Benavent, Martínez, & Flández, 2017)	N=133, NR (6-11; NR)	FMS	Hop, Horizontal jump, Leap, Skip, Side slide, Batting, Catch, Kick, Side arm strike, Stationary dribble, Overhand throw	Composite decision trees used for each skill resulting in classification of (1) Mature (2) Elementary (3) Initial
FITNESSGRAM *USA* (Patterson, Bennington and La-Rosa, 2001)	N=84, 57% girls (10-12; NR)	Fitness assessment	PACER, One-Mile Run, Walk Test, Body Fat Percentage (Skinfold and Bioelectrical Impedance Analyzer (BIA), Body Mass Index, Curl-Up, Trunk Lift, 90° Push-Up, Modified Pull-Up, Flexed Arm Hang, Flexibility , Back-Saver Sit and Reach, Shoulder Stretch, flexibility and PA behaviour	Individual scores for each assessment then converted to FITnessGram® classifies fitness levels using discrete zones to allow for more personalized feedback.

Golf Swing and Putt skill Assessment *Aus* *(Barnett, Hardy, Brian & Robertson, 201*	N=43, NR (6-10; 7.8)	FMS	Skill Materials Directions Golf Swing Performance Criteria	Scores for both skills were summed for each child resulting in a potential score range of 0-24
High/Scope beat competence analysis test *Greece* *(Derri et al., 2001)*	N=77, 47% girls (5-7; 6.1)	Rhythm	Four test items: 1) patting knees with both hands at the same time in seated position 2) patting knees with alternating hands in seated position 3) walking in place from standing position 4) toe-tapping pad with alternating feet. The students are required to synchronize the a aforementioned tasks to the steady beat of two musical selections that are comprised of different tempos: (a) 132 beats/minute and (b) 120 beats/minute.	A three-point scale (0-2) is applied for the evaluation
Körperkoordinat ionstest für Kinder (KTK) *Belgium* *(Vandorpe et al., 2011)*	N=2470, 47% girls, (6-12; NR)	Motor skills	Body control Walking backwards Hopping for height Jumping sideways Moving sideways	The raw test scores from each of the four tests can be transformed into motor quotients
MOBAK-3 test *Germany* *(Hermann, Gerlach & Seelig, 2015; Hermann & Seelig, 2017)*	N=317, 55% girls (NR;7.04)	Motor skill	10 test items:Throwing/ throwing and catching, Bouncing, dribbling, balancing, rolling, rope skipping and moving variably	Test items are dichotomously scaled (0 =failed, 1 = passed, both attempts passed = 2 points)

Movement assessment battery for children-2 *Spain* (*López et al., 2011*)	N=119, 48% girls (8-12;10.4)	Motor skill	The three broad motor skill categories that are assessed are Manual Dexterity, Aiming and Catching, and Balance.	Item performance may be a number of points, a number of performance correct or number of errors performed, and number of seconds to complete task.
MUGI *Sweden* (*Ericsson, 2007; 2008*)	N=25, NR (6-7;NR)	Motor skill	9 gross motor tasks measuring two components of motor skills; Balance/bilateral coordination Hand eye coordination	Three levels are used for evaluation of motor skills 0, 1 and 2.
Obstacle Polygon *Croatia* (*Lovrić, Jelaska and Bilić, 2015*)	N=95, 49% girls (NR;8.1)	Motor skill	Space covering skills Resistance overcoming skills Object control skills	The result of the test is the time needed to successfully accomplish four of the tasks
PARAGON *USA* (*Myers & Well, 2015*)	N=65, 59% girls (5-9; NR)	Gardening movements	Gardening motions (bending, carrying, lifting, stretching, watering)	For each time interval the observer chooses 1 of the 7 PA codes and 1 of the 9 garden tasks.
PLAYfun *Canada* (*Cairney, Veldhuizen, Graham, Rodriguez, Bedard, Bremer & Kriellaars, 2018*)	N=215, 48% girls (7-14; NR)	PL	18 different movement tasks within five domains that assess different aspects of a child's movement skills. The five domains are as follows: 1)running 2)locomotor 3) object control—upper body 4) object control—lower body 5) balance, stability, and body control	Children are assessed using a VAS that is 100 mm in length and divided into four categories:

Slalom Movement Test *Bulgaria* (*Bachev and Zlatev, 2014*)	N=99, 41% girls (11-12; NR)	Agility	N/A	Time taken to complete the course
Star Excursion Balance Test *Spain* (*Calatayud et al., 2014*)	N=24 50% girls (10-12; 11)	Balance	N/A	The point at which the participant touched the line was marked by the examiner and measured manually using a measuring tape.
TGMD-3 *USA* (*Ulrich, 2013*)	N=1460, 50% girls (5-10;8.4)	FMS	The TGMD-3 assesses 13 fundamental motor skills, subdivided into two subscales: Locomotor: run, gallop, hop, leap, horizontal jump, slide Ball Skills: two-handed strike, stationary dribble, catch, kick, overhand throw, underhand roll	Each skill is evaluated on three to five performance criteria, 2- trials summed per skill 0 = if a criterion was not performed 1 = if a criterion was performed
The Leger 20m Shuttle Run test *Canada* (*Cairney et al., 2006*)	N=590; NR (9-11;NR)	Fitness assessment	N/A	Score is the level and number of shuttles reached before missing a beep.
Y Balance Test *USA* (*Faigenbaum, Bagley, Boise, Farrell, Bates & Myer, 2015*)	N=188, NR (6.9-12.1; NR)	Balance	N/A	A total composite score was based on the sum of performance in three directions on both legs

ALPHA Fitness Battery (ALPHA); Athletic Skills Track (AST) ½; Bruininks–Oseretsky Test of Motor Proficiency (BOTMP-SF); Canadian Agility and Movement Skills Assessment (CAMSA); EUROFIT; FITNESSGRAM (FG); FG-COMPASS (FGCOMP); Golf Swing and Putt skill Assessment (GSPA); High/Scope beat competence analysis test (HS); Körperkoordinationstest für Kinder (KTK); Motorische Basiskompetenzen in der 3 (MOBAK-3); Movement assessment battery for children-2 (MABC2); Motorisk Utveckling som Grund för Inlärning (MUGI); Obstacle Polygon (OP); PA Research and Assessment tool for Garden Observation (PARAGON); PlayFUN (PF); Slalom Movement Test (SMT); Star Excursion Balance Test (SEBT); Test of Gross Motor Development-3 (TGMD-3); The Leger 20m Shuttle Run test (20MSR); Y Balance Test (YBT))

PL alignment of physical assessments

Each tool within the review assessed an element of PL and as highlighted in table 4.4, 21 different tools assessed elements within the physical domain. While all included tools assessed an aspect of movement skills on land; no tool considered movement skills in water. Additionally, fundamental movement skills were well represented within the review with 12 tools assessing object control (57%), 15 tools including the elements of locomotor skills and balance (71%). The elements of cardiorespiratory fitness (n=6; 29%), muscular endurance (n=7; 33%), coordination (n=15; 71%), flexibility (n=5; 24%) and agility (n=7; 33%) were represented within existing tools. However, there was a considerable lack of tools available to assess the elements of rhythm (n=1; 5%), speed (n=3; 14%) and sequencing (n=2; 10%), with no tools assessing the elements of progression and an application of movement specific to environment. Finally, the assessment tools within the review that included the most extensive range of PL elements were CAPL-2 (n=10; 48%; object control, locomotor skills, balance, movement skills on land, cardiovascular endurance, muscular endurance, co-ordination, agility, strength and sequencing) and MOBAK-3 (n=10; 48%; object control, locomotor skills, balance, movement skills on land, cardiovascular endurance, co-ordination, flexibility, agility, strength and sequencing).

Table 4.4 An overview of the alignment of each assessment to the physical competence
domain

Assessment Tool	Object control	Balance	Locomotor	M.S (land)	M.S (water)	Moving using equipment	Cardiovascular endurance	Muscular endurance	Coordination	Flexibility	Agility	Strength	Reaction Time	Speed	Power	Rhythmic	Aesthetic/ expressive	Sequencing	Specific to an environment	Progression
ALPHA				•			•	•	•		•	•		•						
AST	•	•	•	•					•											
BOTMP	•	•	•	•				•	•		•	•		•						
CAPL-2	•	•	•	•				•	•		•	•							•	
EUROFIT		•		•				•	•	•		•		•						
FG			•	•				•	•	•		•								
FGCOMP	•	•	•	•				•	•											
GSPA	•			•					•											
HS			•	•					•								•			
KTK		•	•	•					•			•								
20MST			•	•			•													
MACB 2	•	•	•	•					•											
MOBAK 3	•	•	•	•		•			•	•	•								•	
MUGI	•	•	•	•					•								•			

OP   • • • •   •   • • •

PARAGON   • •   •   •   •

PF   • • •   •   •

SMT   • •   •   •

SEBT   • •

TGMD-3   • • • •   •   •

YBT   • •   •

ALPHA Fitness Battery (ALPHA); Athletic Skills Track (AST) ½; Bruininks–Oseretsky Test of Motor Proficiency (BOTMP-SF); Canadian Agility and Movement Skills Assessment (CAMSA); EUROFIT; FITNESSGRAM (FG); FG-COMPASS (FGCOMP); Golf Swing and Putt skill Assessment (GSPA); High/Scope beat competence analysis test (HS); Körperkoordinationstest für Kinder (KTK); Motorische Basiskompetenzen in der 3 (MOBAK-3); Movement assessment battery for children-2 (MABC2); Motorisk Utveckling som Grund för Inlärning (MUGI); Obstacle Polygon (OP); PA Research and Assessment tool for Garden Observation (PARAGON); PlayFUN (PF); Slalom Movement Test (SMT); Star Excursion Balance Test (SEBT); Test of Gross Motor Development-3 (TGMD-3); The Leger 20m Shuttle Run test (20MSR); Y Balance Test (YBT)

Psychometric properties of physical assessments

Table 4.5 displays the results of the psychometric properties of assessments within the physical domain. The results of the review found that within the physical domain a majority of tools reported psychometric qualities relating to reliability and typically at an adequate level of detail, i.e. 82% scored adequate for reliability testing. Further, as highlighted in table 4.5, PROM development (adequate, 26%) and content validity (not reported, 91%), were found to be reported inconsistently across all studies. Similarly, other aspects of validity were reported infrequently, i.e. structural validity (not reported, 48%), cross-cultural validity (not reported, 61%) and criterion validity (not reported, 48%). Notably, seven tools (33%); CAMSA, BOTMP-SF, TGMD-3, GSPA, FGRAM, KTK, MOBAK received a score of 'adequate' for content validity with CAMSA, FGRAM and MOBAK also scoring 'adequate' for content validity.

Table 4.5 An overview of the quality appraisal scores for each assessment within the physical domain

Assessment Tool	PROM development	Content validity	Internal consistency	Reliability	Measurement error	Structural validity	Hypothesis testing	Cross cultural validity	Criterion validity	Responsiveness
ALPHA	IN	NR	NR	IN	IN	NR	NR	NR	NR	NR
APM	NR	NR	NR	A	NR	NR	NR	NR	NR	NR
AST	IN	NR	NR	A	NR	NR	NR	NR	IN	D
BOTMP-SF	A	D	VG	A	IN	A	NR	NR	D	NR
CAMSA	A	A	NR	A	NR	NR	NR	IN	IN	IN
EUROFIT	D	NR	IN	D	VG	VG	IN	IN	IN	IN
FGCOMP	D	NR	IN	A	NR	NR	NR	NR	NR	NR
FGRAM	A	A	A	A	D	D	NR	D	D	NR
GSPA	A	NR	NR	A	NR	IN	IN	IN	IN	IN
HS	NR	NR	VG	A	NR	NR	IN	IN	NR	NR
KTK	A	NR	IN	A	IN	VG	IN	IN	IN	NR
20MS	NR	NR	NR	A	NR	NR	NR	NR	A	NR
MABC-2	D	NR	NR	A	NR	IN	NR	IN	IN	NR
MOBAK	A	A	NR	IN	NR	A	NR	NR	NR	NR
MUGI	IN	NR	D	A	NR	A	NR	NR	NR	IN
OP	D	NR	D	A	NR	IN	IN	IN	VG	NR
OTGAM	D	NR	NR	A	NR	NR	NR	NR	IN	NR
PARAGON	NR	N	NR	A	NR	NR	NR	NR	NR	NR
PF	NR	NR	A	A	NR	A	NR	NR	NR	NR
SMT	NR	NR	NR	D	IN	NR	NR	NR	NR	IN
SEBT	NR	NR	NR	A	A	IN	IN	NR	NR	NR
TGMD-3	A	N	VG	A	NR	VG	NR	VG	IN	NR
YBT	NR	NR	NR	A	IN	NR	NR	NR	NR	NR

(NR= not reported, IN=inadequate, D= doubtful, A=adequate, VG= very goo*d. AMP-inventory manual (AMP);ALPHA Fitness Battery (ALPHA); Athletic Skills Track (AST) ½; Bruininks–Oseretsky Test of Motor Proficiency (BOTMP-SF); Canadian Agility and Movement Skills Assessment (CAMSA); EUROFIT; FITNESSGRAM (FG); FG-COMPASS (FGCOMP); Golf Swing and Putt skill Assessment (GSPA); High/Scope beat competence analysis test (HS); Körperkoordinationstest für Kinder (KTK); Motorische Basiskompetenzen in der 3 (MOBAK-3); Movement assessment battery for children-2 (MABC2); Motorisk Utveckling som Grund för Inlärning (MUGI); Obstacle Polygon (OP); Observation tool for active gaming (OTGM); Physical Activity Research and Assessment tool for Garden Observation (PARAGON); PlayFUN (PF); Slalom Movement Test (SMT); Star Excursion Balance Test (SEBT); Test of Gross Motor Development-3 (TGMD-3); The Leger 20m Shuttle Run test (20MSR); Y Balance Test (YBT)*

Feasibility of physical assessments

Table 4.7 provides the utility matrix ratings of each assessment within the physical domain. The findings related to feasibility highlight that typically a short time was required to complete an assessment (30% <15 min, 35% <30 min). Additionally, the equipment needed to conduct assessments was scored positively for the majority of tools, with equipment likely to be present in a typical primary school (26%) or only some extra equipment required (44%). Furthermore, while the majority of assessments required either a P.E/Sports specialist/researcher to administer (78%), it was also found that typically little or no additional training was required (65%). Across the physical domain, participant understanding and the number of incomplete assessments was not reported consistently. Specifically, within the 21 different physical assessment only two tools (10%) checked participant understanding as part of the implementation instructions (MUGI and PLAYfun) while only three assessments (14%) (CAMSA, EUROFIT and FGram) reported the prevalence of incomplete assessments.

Table 4.7 Feasibility scores for each assessment tool within the physical domain

Assessment Tool	Time	Space	Equipment	Qualification	Training	Participant understanding	Incomplete assessments
ALPHA	**	*	**	**	***	NR	NR
AST	***	**	***	**	***	NR	NR
BOTMP-SF	***	**	*	*	*	NR	NR
CAMSA	****	**	****	**	***	NR	****
EUROFIT	**	*	***	**	***	NR	***
FG	*	*	****	**	***	NR	**
FGCOMP	***	**	***	**	***	NR	NR
GSPA	****	*	*	**	*	NR	NR
HS	***	****	***	****	****	NR	NR
KTK	***	**	**	**	**	NR	NR
20MSR	****	*	***	***	****	NR	NR
MABC2	**	***	**	**	*	NR	NR
MOBAK	***	**	***	**	***	NR	NR
MUGI	**	**	***	**	**	**	NR
OP	****	***	***	**	**	NR	NR
PARAGON	**	**	*	***	**	NR	NR
PF	***	***	***	**	***	****	NR
SMT	****	***	****	**	***	NR	NR
SEBT	****	****	****	**	***	NR	NR
TGMD-3	**	***	****	**	**	NR	NR
YBT	****	****	****	**	***	NR	NR

****excellent, ***good, **fair, *poor

ALPHA Fitness Battery (ALPHA); Athletic Skills Track (AST) ½; Bruininks–Oseretsky Test of Motor Proficiency (BOTMP-SF); Canadian Agility and Movement Skills Assessment (CAMSA); EUROFIT; FITNESSGRAM (FG); FG-COMPASS (FGCOMP); Golf Swing and Putt skill Assessment (GSPA); High/Scope beat competence analysis test (HS); Körperkoordinationstest für Kinder (KTK); Motorische Basiskompetenzen in der 3 (MOBAK-3); Movement assessment battery for children-2 (MABC2); Motorisk Utveckling som Grund för Inlärning (MUGI); Obstacle Polygon (OP); PA Research and Assessment tool for Garden Observation (PARAGON); PlayFUN (PF); Slalom Movement Test (SMT); Star Excursion Balance Test (SEBT); Test of Gross Motor Development-3 (TGMD-3); The Leger 20m Shuttle Run test (20MSR); Y Balance Test (YBT)

4.6 Results of the Affective Domain

The results of the affective domain include 60 studies providing information for 33 assessment tools within the 7-11 age range. Table 4.8 includes the study characterises and details information relating to geographical location, setting, age range and scoring each relate to study characteristics. The assessments were conducted within the USA (n=23), UK (n=4), Australia (n=3), Canada (n=1), Hong Kong (n=1) and Turkey (n=1). Assessments were typically administered within the school setting using questionnaires (n=26), surveys (n=3), self-report inventory (n=1), interview (n=1), observation (n=1) and ecological momentary assessment (n=1). Both sample size and age range were well reported with the average sample size of (n=494) and mean age of (8.59 years).

Table 4.8 The study characteristic information for each assessment within the affective domain

Assessment and country of origin	Participant n, gender (%) (age range; mean age)	Purpose/ use of assessment	Mode of Assessment	Scale Design	Scale Scoring
Trichotomous Achievement Goal Model *Turkey* *(Agbuga, 2009)*	N=158, 57% female, (8-12, NR)	Assess trichotomous achievement goal theory in elementary PE	Questionnaire	15 items reflecting mastery, performance approach and performance avoidance achievement goals. Each item prefaced 'in my PE classes…'	5-point Likert scale (not at all true to very true)
FHC-Q *USA* *(Bandelli et al., 2016)*	N=284, 49.3% female (9-12, NR)	Assess energy related behaviours including intake of fruits and vegetables, sugar-sweetened beverages, processed packaged snacks, and fast food; PA; recreational screen time; and associated psychosocial determinants	Questionnaire. Utilised Audience Response System through PowerPoint.	71 items in total: Self-determination (9 questions), Outcome expectations (15 questions), Self-efficacy (20) questions), Habit strength (6 questions), Goal intention (6 questions), Knowledge (6 questions). Social desirability (9 questions).	5-point Likert scale (options varied)
ATOP Scales *USA* *(Beyer, 2015)*	N= 362, 49% female (9-13, 11)	Quantitatively assess attitudes towards outdoor play	Ipad based questionnaire	Three scales: Perceived benefit of playing outside 4 items; Extent to which students enjoy unstructured play 3 items; Barriers to outdoor play 5 items.	'How much do you agree with each statement?' Responses on a 5 point likert scale from strongly disagree to strongly agree

ASK-KIDS Inventory *Australia* *(Bornholt & Piccolo, 2005)*	N=76, 43% female (4-13.5, 8.1±2.3)	Assess self-concept in relation to physical movement, natural talent, effort, difficulty, personal identity and social identity.	Self-report inventory	Dot-point rating scores 1 (low) to 5 (high)	Scores averaged from (1) low to (5) high
Feelings about physical movement *Australia* *(Bornholt & Piccolo, 2005)*	N=56 , 43% female (4-11, 8.0±2.1)	Assess feelings about physical movements	One to one interview	Diagram (stick figures running and catching) researcher reads accompanying paragraph and the child ticks as many words as needed in relation to five general feelings	Responses scaled from 1 (low) to 7 (high)
AGSYS *USA* *(Cumming et al., 2008)*	N=1675 NR (9-12, NR)	Use the 2x2 achievement goal framework to assess goal approach orientations	Questionnaire	12 items related to mastery/ego X approach/avoidance goal framework	5point Likert scale from 1 (not at all true) to 5 (very true)
CPAS *USA* *(DeBate, Huberty & Pettee , 2009)*	N=932 100% female (9-14, NR)	Asses PA commitment	Questionnaire	12 items measuring attitudes and feelings towards PA	Likert scale 0 (strongly disagree) to 3 (strongly agree)
PA Beliefs and Motives *USA* *(Dishman,*	N=2092, 4853 female (10-12, NR)	Assess motives for PA	Questionnaire	Self-efficacy (8 items). Perceived barriers: 3 scales; obstacles (3 items), evaluation (3 items), outcomes (3 items). Motives for PA: 30 items, 5 scales for intrinsic; enjoyment (7 items), competence (7 items) and extrinsic; fitness (5	All used 4 point order response format apart from perceived parental support, 5 point ordered format

Saunders, McIver, Dowda and Pate, 2013)				items), appearance (6 items), social (5 items). Parental support (5 items).	
Momentary Assessment of Affect and Feeling States *USA* *(Dunton et al., 2014)*	N=119, 48% female (9-13, NR)	Use Ecological Momentary Assessment (EMA) to bi-directionally examine affective and feeling states relate to PA	EMA	Positive affect, negative affect, physical feeling states all assessed by 2 items each when prompted through a mobile phone	Response options included 0=not at all, 1=a little bit, 2=quite a bit, 3=extremely
LEAP *Australia* *(Hyndman, Telford, Finch, Ullah and Benson 2013; 2014)*	N=197, 43% female (8-12, NR)	Assess children's enjoyment of lunchtime play	Survey	Children completed 'expected' (before lunch) and 'actual' (after lunch) enjoyment of lunch time play using survey cards with pictorial scale	5 point Likert pictorial scale from very unhappy (1) to very happy (5)
ATCPE *UK* *(Jones, 1988)*	N=223 NR (9-12, NR)	Assess attitudes towards curriculum PE	Questionnaire	25 items (13 positive and 12 negative)	5 point Likert scale from strongly disagree (1) to strongly agree (5)
RCS *USA* *(Lakes, 2013)*	N=112, 51% female (NR, 4-11)	Assess children's self-regulatory abilities in physically active context	Observation of PA challenge course	16 items and three subscales: Cognitive Self-Regulation (6 items, including "control over emotions- uncontrolled emotions")	Bipolar adjectives (e.g., "attentive – inattentive") are used for each item, and raters were asked to rate the child using a 7-point scale.

Self-Efficacy Scale *USA* *(Leary, Ice and Cottrell, 2012)*	N=15, 68% female (NR, 8.2±0.85)	Assess children's self-efficacy for overcoming PA barriers	Survey	11 items starting with the stem 'Can you…'	5 point Likert scale
Questionnaire to measure PA self-efficacy, enjoyment, social support *Hong Kong* *(Liang, Lau, Huang, Maddison and Baranowski 2014)*	N= 457, 50% female (8-12, 10.3±1.0))	Assess PA self-efficacy, enjoyment, social support among Hong Kong Chinese children	Questionnaire	8 item scale used to measure PA self-efficacy. 7 item scale to assess PA enjoyment. 10 items to assess social support for exercise	Self-efficacy and enjoyment scales used Likert scale ranging from 1 (Disagree a lot) to 5 (Agree a lot). Social support scale used a 5 point scale 1 (none) to 5 (very often)
CAPL-2 *Canada* *(Tremblay et al., 2018)*	N=1034, 50.1% girls (8-10; 10.1±1.17)	PL	Questionnaire	12 multiple choice questions: 'What's most like me?' (6 items) 'Why are you active?' (3 items); 'How do you feel about being active?' (3 items)	5 point Likert scale 1(not true for me) to 5 (very true for me)
PACES *USA* *(Moore, Yin, Hanes, Duda, Gutin and Barbeau 2009)*	N=564 53% female (8-9; 8.72±.54)	Assess the enjoyment of PA	Questionnaire	16 bipolar statements starting with the stem 'When I am physically active…'	5 point Likert scale 1 (Disagree a lot) to 5 (Agree a lot)

CY-PSPP *USA* *(Welk, Corbin, Dowell and Harris, 1997)*	N=152 53% female (9-11; NR)	Assess physical self-perceptions in children	Questionnaire	36 items, 6 items for each of the 6 domains (global self-esteem, physical self-worth, sport competence, body attractiveness, physical strength, physical condition)	4 point structured alternate format and standard 4 point Likert scale for comparison
PAHFE *USA* *(Perry et al., 2008)*	N=131 54% female (8-14; 9.89±.NR)	Assess personal goal setting and decision making efficacy for PA and food choices	Questionnaire	18 items representing children may experiences when attempting to improve PA and eating behaviours	5 point Likert scale from 1 (not sure at all) to 5 (completely sure)
MOSS *USA* *(Weiss, Bredemeier and Shewchuk, 1985; Allen, 2005)*	N=155 45% female (8-12; 10.2±1.4)	Assess children's motivational orientation for engagement in PA	Questionnaire	27 items, 5 subscales: Challenge (5 items relating to preference for challenging or easy skills), curiosity (4 items relating to desire to participate), mastery (5 items relating to problem solving and mastery attempts), judgement (6 items relating to self-assessment vs teacher assessment), criteria (7 items relating to preference for internal sense of success/failure vs external determined success/failure)	Structured alternative scoring 1 (low) to 4 (high). Children indicate if 'Sort of true for me' or 'really true for me'. Separate scores given for each subscale. High scores indicate more intrinsic motivation
CAPA *USA* *(Brustad 1993; 2009)*	N=81 53% female (9-10; 10.4±0.3)	Measurement of attraction to PA	Questionnaire. Adapted version used individual verbal questioning	Original scale has 25 items (5 subscales with 5 items each), shorter scale has 15 items (5 subscales with 3 items each). Subscales include: liking of games and sports, liking of physical exertion and exercise, liking of vigorous PA, peer acceptance in sport and games, importance of exercise	Structured alternate. Adapted version used 1 to 4 Likert scale
HOP'N Evaluation	N=230 51% female	Assess psychosocial variables as part of a 3 year randomised	Questionnaire	16 items: PA task self-efficacy (1 item), PA barriers self-efficacy (4 items), PA enjoyment (2 items), Perceived opportunity for PA (2 items),	3 point scale (e.g. not sure at all-somewhat sure- very sure). Perceived habitual PA scores were

USA *(Rosenkranz, Welk, Hastman and Dzewaaltowski, 2011)*	(9-10; 9.5±0.7)	controlled trial aiming to prevent obesity through an after-school program		Perceived habitual PA (2 items), and perceived parental support (5 items)	assessed using a 2 item screener, averaged and dichotomized a meeting PA guidelines or not. Parental support was rated on a 6 point scale (never to daily)
PASES *USA* *(Bartholomew et al., 2006)*	N=442 NR (10-11; NR)	Assess psychosocial determinants on children's PA: social influences, self-efficacy, beliefs and intention	Questionnaire	Social influences (1 factor), self-efficacy (3 factors; support seeking, barriers, positive alternatives), beliefs (2 factors; social outcomes, PA outcomes)	2 point scale (yes or no)
PA Motivation Scales UK *(Sebire, Jago, Fox, Edwards and Thompson, 2013)*	N=462 56.9% female (7-11; 10.03±0.566)	Assess self-determined motivation for PA and PA psychological need satisfaction in children	Questionnaire	Self-determined motivation for PA: 12 items, 3 per motivation scale (intrinsic, extrinsic, external). PA psychological need satisfaction: autonomy (6 items), competence (6 items), relatedness (6 items)	5 point Likert scale from 1 (not true for me) to 5 (very true for me).
SPPC *UK* *(Harter, 1982; Muris, Meesters and Fijen, 2003)*	N=2394, NR (8-12, NR)	Assess perceived competence in children	Questionnaire	36 items, 5 domain specific sub-scale each with 6 items: scholastic competence, social acceptance, athletic competence, physical appearance, behavioural conduct. One global measure of self-worth	Structure alternative format
Enjoyment in PE	N=148 47%	Assess students enjoyment in PE and	Survey	10 statements relating to enjoyment (7) and	5 point Likert strongly disagree (1)

USA	female	exergaming		perceived exertion (3)	strongly agree (5)
(Shewmake et al., 2015)	(8-10; NR)				
CATPA *USA* *(Simon and Smoll, 1974; 1981)*	N=992 51% female (9-12; NR)	Asses attitudes towards PA	Questionnaire	6 scales; social, health and fitness, pursuit of vertigo, aesthetic, catharsis and aesthetic. Each had 8 items	5 point Likert scale and semantic differential technique with a 0-7 bipolar continuum, with 0 as a neutral reference point. Adjectives at each end of the continuum included; good-bad, of no use-useful, not pleasant-pleasant, bitter-sweet, nice-awful, happy-sad, dirty-clean, steady-nervous
Negative Attitudes towards PA Scale *USA* *(Nelson, Benson, Jenson, 2000)*	N=382 =46% female (10-12; 10.8±0.65)	Measure negative attitudes towards PA	Questionnaire	All items (9) followed the stem 'If I were to be physically active on most days…'	5 point Likert scale from 1 (disagree a lot) to 5 (agree a lot)
Positive Attitudes towards PA Scale *USA* *(Nelson, Benson, Jenson, 2000)*	N=382 =46% female (10-12; 10.8±0.65)	Measure positive attitudes towards PA	Questionnaire	All items (8) followed the stem 'If I were to be physically active on most days…'	5 point Likert scale from 1 (disagree a lot) to 5 (agree a lot)
PLOC in PE *UK*	N=817 50% female	Assess the revised PLOC for use in PE	Questionnaire	PLOC scale adapted for PE (19 items), perceived autonomy support (6 items), subjective vitality (unclear how many items)	Participants provided their responses on a 1-5 Likert type scale anchored by 1 (totally disagree) 4

(Vlachopoulos, Katarti, Kontou, Moustaka and Goudas, 2011)	(11-12; NR)					(in between) and 7 (totally agree)
DPAPI *USA* *(Chen, 2011)*	N=435 51% female (11-12; 9.87±1.14)	Assess psychological needs, motivational types and motivational consequences for PA participation outside of school	Questionnaire	Innate psychological needs (6 items), motivational types (12 items), motivational consequences (6 items)		Innate psychological needs, motivational types and motivational behavioural consequences were assessed on a 5 point Likert scale 5 (very like me) to 1 (not like me). Responses to motivational affective consequences included 4 semantic pairs anchored on a 5 point Likert scale with smiley faces
PMSQ *USA* *(Xiang, Bruene, & McBride, 2004)*	N=116 42% female (9-10; NR)	Assess perceptions of the motivational climate of team in terms of matter and performance goals	Questionnaire	Statement starts with stem 'In roadrunners…' followed by 24 items related to perception of motivational climate. 11 mastery focussed and 13 performance focussed items. In original scale (used with older children) 9 mastery and 12 performance items related team…'		Participants responded in agreement to statements on a 5 point Likert scales from YES (5) to NO (1) (YES, yes ?, no, NO), scores calculated by an average for each scale.. In original scale 1 (strongly disagree) to 5 (strongly agree)

Not Reported (NR); PA (PA); PE (PE); United Kingdom (UK); United States of America (USA); Food, Health, and Choices Questionnaire (FHC-Q); Attitudes Towards Outdoor Play Scales (ATOP); Achievement Goal Scale for Youth Sports (AGSYS); Commitment to PA Scale (CPAS); Lunchtime Enjoyment Activity and Play (LEAP); PA Self-Efficacy Questionnaire (PASE): Attitudes Towards Curriculum PE Scale (ATCPE); Response to Challenge Scale (RCS); Canadian Assessment of PL (CAPL-2); PA Enjoyment Scale (PACES); The Children and Youth Self-Perception Profile (CY-PSPP); Children's Perception of Motor Competence Scale (CMPI);PA and Healthy Food Efficacy (PAHFE); Motivational Orientation in Sport Questionnaire (MOSS);Children's Attraction to PA (CAPA); Healthy Opportunities for PA and Nutrition (HOP'N Evaluation); PA Self- Efficacy Scales (PASES); Self-Perception Profile for Children (SPPC); Children's Attitudes toward PA (CATPA); Determinants of PA Participatory Inventory (DPAPI): Perceived Motivational Climate in Sport Questionnaire (PMCSQ)

PL alignment of affective assessments

Table 4.9 displays the PL alignment results for assessments within the affective domain. Each affective tool within the review assessed at least one element of PL and many of these related directly to the affective domain, as highlighted in table 4.9. Specifically, within the assessments that were appropriate for use within children aged 7-11 years, 33(56%) affective assessments assessed elements within the affective domain of PL. Within this domain, 11(33%) of tools assessed an aspect of motivation making it the most frequently assessed element. In addition, 27% of assessments also included measure relating to confidence, while three assessment tools (9%) considered both confidence and motivation together within the same assessment: Food, Health, and Choices Questionnaire (FHC-Q), PA Beliefs and Motives (PABM) and the Canadian Assessment for Physical Literacy-2 (CAPL-2). In addition, 27% of assessments also included a measure relating to self-perception and perceived competence, while 24% of assessments included an assessment related to an individual's comfortability and connectedness to the world. The elements of emotional regulation (18%), children's Attitudes toward PA, resilience (15%) and adaptability (9%) were reported less frequently across the affective domain. Notably, no assessments included a measure of empathy. The assessment tool within the review that included the most extensive range of PL elements within the affective domain was PABM (n=5; 15%; motivation, confidence, enjoyment/happiness, adaptability, willingness to try new things).

Table 4.9 An overview of the alignment of each assessment to the affective domain.

Assessment Tool	Confidence	Motivation	Emotional regulation	Enjoyment/ happiness	Empathy	Persistence/resilience/ commitment	Adaptability	Willingness to try new activities	Autonomy	Comfortable and connected with the world	Self-perception/ self-esteem	Perceived physical competence
Trichotomous		•										
FHC-Q	•	•										
PMSC												•
ATOP Scale			•									
ASK-KIDS						•				•	•	
FAPM			•									
AGYS		•				•						
CPAS				•			•	•		•		

PABM	•	•		•	•		•		
MAAP				•					
LEAP				•					
PASE	•						•	•	•
ATCPE			•	•			•		•
RCS			•						
Self-Efficacy	•								
QPA	•			•		•			
CAPL-2	•	•		•					•
PACES				•					
CY-PSSP								•	•
TEOSQ		•							
CMPI									•
PAHFE	•								

MOSS

CAPA

HOP'N

PASE

PAMS

SPPC

EnjoyPE

CATPA

NAS

PAS

PLOC in PE

DPAPI

PMSQ

PA (PA); PE (PE); Food, Health, and Choices Questionnaire (FHC-Q); Pictorial Scale for Perceived Movement Skill Competence for young children (PMSC); Attitudes Towards Outdoor Play Scales (ATOP); Achievement Goal Scale for Youth Sports (AGSYS); Commitment to PA Scale (CPAS); Lunchtime Enjoyment Activity and Play (LEAP) PA Self-Efficacy Questionnaire (PASE); Attitudes Towards Curriculum PE Scale (ATCPE); Response to Challenge Scale (RCS); Canadian Assessment of PL (CAPL-2) PA Enjoyment Scale (PACES); The Children and Youth Self-Perception Profile (CY-PSPP); Task and Ego Orientation in Sport Questionnaire (TEOSQ); Children's Perception of Motor Competence Scale (CMPI); PA and Healthy Food Efficacy (PAHFE); Motivational Orientation in Sport Questionnaire (MOSS); Children's Attraction to PA (CAPA); Healthy Opportunities for PA and Nutrition (HOP'N Evaluation); PA Self- Efficacy Scales (PASES); Self-Perception Profile for Children (SPPC); Children's Attitudes toward PA (CATPA); Perceived Locus of Causality in PE (PLOC in PE); Determinants of PA Participatory Inventory (DPAPI): Perceived Motivational Climate in Sport Questionnaire (PMCSQ)

Psychometric properties of affective assessments

Psychometric properties of included affective assessments are displayed in Table 4.10. The majority of tools reported psychometric qualities relating to reliability and typically at an adequate level of detail, i.e. 88% scored 'very good' for internal consistency testing. While structural validity was reported frequently and to a high standard, (very good, 46%), cross-cultural validity was typically not reported (not reported, 93%), while responsiveness, criterion validity and measurement error were also typically not reported. In addition, both PROM development and content validity were not reported consistently across studies, with 67% of studies scoring 'inadequate' for both properties.

Feasibility of affective assessments

Feasibility of affective assessments are displayed in Table 4.11. Typically, the time to complete each assessment was not reported (42%). For assessments that did report timing information these were scored poorly, taking between 30-60 minutes to complete (24%). The majority of assessments (94%) used equipment likely to be present in a typical primary/elementary school setting or only some extra equipment required. It was generally not reported if qualifications (not reported, 84%) or further training (not reported, 87%) was required to conduct assessments. For some assessments, participant understanding was either investigated (33%) or estimated (6%).

Table 4.10 An overview of the quality appraisal scores for each assessment within the affective domain

Assessment Tool	PROM Develop	Content Validity	Structural validity	Internal consistency	Cross cultural validity	Reliability	Measurement error	Criterion validity	Hypothesis testing	Responsiveness
Trichotomous	IN	D	VG	VG	NR	NR	NR	NR	NR	NR
FHC-Q	A	A	IN	VG	NR	VG	NR	NR	NR	NR
ATOP	D	D	A	VG	NR	NR	NR	NR	NR	NR
ASK-KIDS	IN	IN	D	IN	NR	NR	NR	NR	MR	NR
FAPM	IN	IN	NR	VG	NR	NR	NR	NR	NR	NR
AGSYS	D	IN	VG	VG	NR	A	NR	NR	A	NR
CPAS	IN	IN	A	VG	NR	NR	NR	NR	NR	NR
PABM	IN	IN	VG	NR	NR	NR	NR	NR	NR	NR
MAAP	IN	IN	NR	VG	NR	NR	NR	NR	NR	NR
LEAP	D	D	NR	NR	NR	VG	NR	NR	NR	NR
PASE	D	IN	A	VG	NR	NR	NR	NR	NR	NR
ATCPE	D	NR	NR	VG	NR	D	NR	NR	NR	NR
RCS	D	D	NR	NR	NR	A	NR	NR	VG	NR
Self-efficacy	D	D	NR	VG	NR	NR	NR	NR	NR	NR
QPA	D	IN	VG	VG	NR	VG	NR	NR	NR	NR
CAPL-2	IN	VG	VG	NR	NR	NR	NR	NR	D	NR
PACES	IN	IN	VG	VG	NR	NR	NR	NR	NR	NR
CY-PSPP	IN	IN	VG	VG	VG	IN	NR	NR	NR	NR
TEOSQ	IN	D	NR	VG	NR	NR	NR	NR	NR	NR
PAHFE	D	D	VG	VG	NR	IN	NR	NR	NR	NR
MOSS	IN	IN	IN	VG	NR	NR	NR	NR	NR	NR
CAPA	D	IN	VG	VG	IN	NR	NR	NR	NR	NR
HOP'N	IN	IN	NR	VG	NR	NR	NR	NR	NR	NR
PASES	D	IN	A	VG	VG	D	NR	NR	NR	NR
PAMS	IN	NR	VG	VG	NR	NR	NR	NR	NR	NR
SPPC	D	IN	IN	VG	NR	A	NR	NR	NR	NR

EnjoyPE	IN	NR	NR	VG	NR	NR	NR	NR	NR	NR
CATPA	NR	IN	NR	VG	NR	A	NR	NR	NR	NR
NAS	IN	IN	VG	VG	NR	NR	NR	NR	NR	NR
PAS	IN	IN	VG	VG	NR	NR	NR	NR	NR	NR
PLOC in PE	IN	NR	VG	VG	NR	NR	NR	NR	NR	NR
DPAPI	IN	IN	VG	NR	NR	NR	NR	NR	NR	NR
PMCS	IN	IN	NR	VG	NR	A	NR	NR	NR	NR

(NR= not reported, IN=inadequate, D= doubtful, A=adequate, VG= very good)

PA (PA); PE (PE); PE (PE); Food, Health, and Choices Questionnaire (FHC-Q); Pictorial Scale for Perceived Movement Skill Competence for young children (PMSC); Attitudes Towards Outdoor Play Scales (ATOP); Achievement Goal Scale for Youth Sports (AGSYS); Commitment to PA Scale (CPAS); Lunchtime Enjoyment Activity and Play (LEAP); PA Self-Efficacy Questionnaire (PASE): Attitudes Towards Curriculum PE Scale (ATCPE); Response to Challenge Scale (RCS); Canadian Assessment of PL (CAPL-2); PA Enjoyment Scale (PACES); The Children and Youth Self-Perception Profile (CY-PSPP); Task and Ego Orientation in Sport Questionnaire (TEOSQ); Children's Perception of Motor Competence Scale (CMPI);Motivational Orientation in Sport Questionnaire (MOSS); Children's Attraction to PA (CAPA); Healthy Opportunities for PA and Nutrition (HOP'N Evaluation); PA Self- Efficacy Scales (PASES); Self-Perception Profile for Children (SPPC); Children's Attitudes toward PA (CATPA); Perceived Locus of Causality in PE (PLOC in PE); Determinants of PA Participatory Inventory (DPAPI):

Table 4.11 Feasibility scores for each assessment related to the affective domain.

Assessment Tool	Time	Space	Equipment	Qualification	Training	Participant understanding	Incomplete assessments
Trichotomous	**	*****	****	NR	NR	***	NR
FHC-Q	**	***	***	NR	NR	NR	NR
ATOP	NR	****	***	NR	NR	****	***
ASK-KIDS	****	****	****	NR	NR	NR	NR
FAPM	NR	***	****	NR	NR	NR	NR
AGSYS	NR	***	****	NR	NR	****	NR
CPAS	**	****	****	NR	*	NR	NR
PABM	NR	****	****	NR	NR	NR	NR
MAAP	*	****	*	NR	NR	NR	**
LEAP	NR	****	****	NR	NR	****	***
PASE	NR	****	***	NR	NR	****	**
ATCPE	****	****	***	NR	NR	****	NR
RCS	NR	**	**	*	***	NR	NR
Self-efficacy	**	****	****	NR	NR	****	NR
QPA	***	****	****	NR	NR	****	NR
CAPL-2	NR	****	****	*	***	***	***
PACES	**	****	****	NR	NR	****	****
CY-PSPP	NR	****	****	NR	NR	NR	NR
TEOSQ	**	****	****	NR	NR	NR	NR

PAHFE	***	****	****	NR	NR	****	NR
MOSS	***	***	****	NR	NR	NR	NR
CAPA	***	***	****	NR	*	****	NR
HOP'N	NR	****	****	NR	NR	NR	NR
PASES	**	****	****	NR	NR	****	NR
PAMS	NR	****	****	NR	NR	NR	NR
SPPC	NR	****	****	*	NR	NR	NR
EnjoyPE	****	****	****	NR	NR	NR	NR
CATPA	NR	****	****	NR	NR	NR	NR
NAS	NR	****	****	*	NR	NR	NR
PAS	NR	****	****	*	NR	NR	NR
PLOC in PE	NR	****	****	NR	NR	NR	NR
DPAPI	NR	****	****	NR	NR	NR	****
PMCS	**	****	****	NR	NR	NR	NR

****excellent, ***good, **fair, *poor

Not reported (NR);PA (PA); PE (PE);Food, Health, and Choices Questionnaire (FHC-Q); Pictorial Scale for Perceived Movement Skill Competence for young children (PMSC); Attitudes Towards Outdoor Play Scales (ATOP); Achievement Goal Scale for Youth Sports (AGSYS); Commitment to PA Scale (CPAS); Lunchtime Enjoyment Activity and Play (LEAP); PA Self-Efficacy Questionnaire (PASE): Attitudes Towards Curriculum PE Scale (ATCPE); Response to Challenge Scale (RCS); Canadian Assessment of PL (CAPL-2); PA Enjoyment Scale (PACES); The Children and Youth Self-Perception Profile (CY-PSPP); Task and Ego Orientation in Sport Questionnaire (TEOSQ); Children's Perception of Motor Competence Scale (CMPI); PA and Healthy Food Efficacy (PAHFE); Motivational Orientation in Sport Questionnaire (MOSS); Children's Attraction to PA (CAPA); Healthy Opportunities for PA and Nutrition (HOP'N Evaluation); PA Self- Efficacy Scales (PASES); Self-Perception Profile for Children (SPPC); Children's Attitudes toward PA (CATPA); Perceived Locus of Causality in PE (PLOC in PE); Determinants of PA Participatory Inventory (DPAPI): Perceived Motivational Climate in Sport Questionnaire (PMCS)

4.7 Results of the Cognitive Domain

The results of the cognitive domain include 7 studies providing information for 6 assessment tools within the 7-11 age range (see Table 4.12). Three assessments were based upon theoretical frameworks; PL, self-regulation, and the transtheoretical model of change. Three studies were conducted in the USA, with the remaining studies from Australia, Greece, Canada, and Spain. Six studies were conducted within a primary school and assessments were administered typically within the classroom setting. Two studies used a questionnaire, two studies utilised a photo pairs activity, one study observed the completion of a PA obstacle course, and one study used an interview with a picture sorting activity relating to weight bearing PA choices. The eligible studies had sample sizes ranging from 41 to 4171. Reporting of ages varied between mean and ranges, and information provided with in included studies did not allow for consistency in reporting in this review.

PL Alignment of cognitive assessments

Six assessments (10%) measured an aspect of the cognitive domain (see Table 4.13). The majority assessed the benefits of PA (83%), the importance of PA (67%) and the effects of PA on the body (67%). In addition, four assessments (CAPL-2, Pre-FPQ, BONES-PAS and Scales to measure knowledge and preference for diet and PA) assessed three elements within one tool (benefits of PA, importance of PA, effects of PA on body). While, 16% of assessments included a measure related to decision making, the ability to reflect and tactics, rules or strategy, no assessments included a measure of the ability to identify/describe movement or creativity/imagination. The CAPL-2 assessed the most cognitive PL aspects (n=7; 67%; Benefits of PA, importance of PA, effects of PA on the body, PA opportunities, sedentary behaviour, ability to reflect and safety considerations).

Table 4.12 The study characteristic information for each assessment within the cognitive domain

Assessment and country of origin	Participant n, gender (%) (age range; mean age)	Purpose/ use of assessment	Mode of Assessment	Scale Design	Scale Scoring
Pre-FPQ *Australia* *(Wiseman, Harris & Downes 2017)*	N=86, 48% female, (NR; 51±7.5 months)	Assess child food and activity knowledge and preferences	Electronic questionnaire	Four subscales: food knowledge (Score out of 10), PA knowledge (8), food preference (10) and PA preference (8)	Sum of healthy choices made (healthful choice= 1-point, unhealthful choice/ sedentary behaviour = 0)
BONES-PAS *USA* *(Economos, Hennessey, Sacheck, Shea & Naumova, 2010)*	N=41, 63% female, NR, 7.1±0.8)	Assess participation in and knowledge of weight-bearing PA	Picture sorting activity	Children given 10 different PA pictures, and 3 coloured placemats with "yes", "no", "I don't know"; "yesterday", "the day before yesterday"; "good for building bones", "not good for building bones", "don't know"	Each correct response scored as 1 and all incorrect scores including "don't know" responses were scored as 0
RCS *USA* *(Lakes, 2013)*	N=207, 51% female (NR, 4-11)	Assess children's self-regulatory abilities in physically active context	Observation of PA challenge course	16 items and three subscales: Cognitive Self-Regulation (7 items, including "attentive – inattentive")	Bipolar adjectives (e.g., "attentive – inattentive") are used for each item, and raters were asked to rate the child using a 7-point scale.
Pupil Health Knowledge Assessment *Greece* *(Manios, Moschandreas, Hatzis, & Kafatos, 1999)*	N= 4171, NR (6-10, NR)	Assess knowledge of diet, food products, and PA before and after 3-year intervention	Questionnaire	Multiple choice questionnaire	NR

CAPL-2 *Canada* *(Tremblay et al., 2018)*	N=10034, 50.1% girls (8-10; 10.1±1.17)	PL	Questionnaire	Multiple choice questions, fill in the gap statements, circle activity and safety equipment	1 point for each correct answer (apart from negatively marked safety question). Total score 18. Scores translated into 'beginning, progressing, achieving and excelling'
Scales to measure knowledge and preference for diet and PA *USA* *(Calfas , Sallis & Nader, 1991)*	N=81 59% female (3-8, NR)	Assess knowledge and preference for diet and PA behaviour	Picture sorting activity	15 photo pairs, including health and unhealthy food (8)/PA (7) (5 of each were retained in final scale)	Responses of 'healthful' or 'unhealthful', then asked to point to the activity they liked best, and rate from 1 (happy face-like a lot) to 3 (sad face-don't like very much)

Not Reported (NR); PA (PA);Preschool PA and Food Questionnaire (Pre-FPQ); Beat Osteoporosis Now-PA Survey (BONES-PAS); Response to Challenge Scale (RCS); Canadian Assessment of PL (CAPL-2); Scales to measure knowledge and preference of diet and PA (SCALES)

Table 4.13 An overview of the alignment of each assessment to the cognitive domain.

Assessment Tool	Benefits of PA	Importance of PA	Effects of PA on body	PA opportunities	Sedentary behaviour	Ability to identify and describe and movement	Creativity and imagination	Decision making	Ability to reflect	Tactics, rules and strategy	Safety considerations and risk
Pre-FPQ	•	•	•		•						
BONES-PAS	•	•	•								
RCS									•		
Pupil Health Knowledge	•										
CAPL-2	•	•	•	•	•				•		•
Scales to measure knowledge and preference for diet and PA	•	•	•								

Psychometric properties of cognitive assessments

Psychometric properties of assessments related to the cognitive domain of PL are appaised in Table 4.14. Within the cognitive domain the majority of tools reported psychometric qualities relating to reliability (83%), however, typically with an 'inadequate' level of detail, i.e. only n=2 (33%) were scored as 'adequate' for reliability and n=1 (16%) were scored as 'very good' for internal consistency testing. Within the cognitive domain CAPL-2, the RCS and PRE-FPQ showed reported the strongest methodological rigour. While CAPL-2 received a score of 'very good' for structural validity; the RCS scored 'adequate' for reliability and Pre-FPQ scored 'very good' for internal consistency and 'adequate' for reliability. Notably, none of the tools within the cognitive domain reported results for responsiveness, criterion validity, cross-cultural validity and measurement error and consequently these psychometric properties could not be appraised.

Table 4.14 An overview of the quality appraisal scores for each assessment within the cognitive domain

Assessment Tool	PROM Develop ment	Content Validity	Internal consistency	Reliability	Measurement error	Structura l validity	Hypothesi s testing	Cross cultural validity	Criterion validity	Respons iveness
Pre-FPQ	D	D	VG	A	NR	NR	NR	NR	NR	NR
BONES	IN	IN	NR	IN	NR	NR	NR	NR	NR	NR
RCS	D	IN	NR	A	NR	NR	VG	NR	NR	NR
PH- KA	IN	IN	NR	IN	NR	NR	NR	NR	NR	NR
CAPL-2	IN	VG	NR	NR	NR	VG	D	NR	NR	NR
SCALES	D	IN	D	IN	NR	NR	NR	NR	NR	NR

(NR= not reported, IN=inadequate, D= doubtful, A=adequate, VG= very good)
Preschool PA and Food Questionnaire (Pre-FPQ); Beat Osteoporosis Now-PA Survey (BONES-PAS); Response to Challenge Scale (RCS); Pupils Health Knowledge Assessment (PH-KA), Canadian Assessment of PL (CAPL-2); Scales to measure knowledge and preference of diet and PA (SCALES)

Feasibility of cognitive assessments

Table 4.15 provides the utility matrix ratings of each assessment. The findings related to feasibility highlight that typically the time to complete each assessment was not reported (67%). However, for assessments that did report information regarding the time to administer, these were typically scored as 'excellent' due to taking less than 5 minutes to complete (33%). In addition, the equipment needed to conduct assessments was typically scored positively for all of the assessments within the cognitive domain (good and excellent, 83%) Despite this, it was not consistently reported whether qualifications (not reported, 50%) or further training (not reported, 50%) were required to conduct assessments. However, for some assessments participant understanding was either investigated (33%) or estimated (16%).

Table 4.15 Feasibility scores for each assessment related to the cognitive domain.

Assessment Tool	Time	Space	Equipment	Qualification	Training	Participant understanding	Incomplete assessments
Pre-FPQ	****	*****	***	NR	NR	****	NR
BONES-PAS	NR	****	***	NR	NR	****	NR
RCS	NR	**	**	*	***	NR	NR
PH-KA	NR	****	***	****	**	NR	*
CAPL-2	NR	****	***	*	***	***	***
SCALES	****	****	***	NR	NR	NR	NR

****excellent, ***good, **fair, *poor
Preschool PA and Food Questionnaire (Pre-FPQ); Beat Osteoporosis Now-PA Survey (BONES-PAS); Response to Challenge Scale (RCS); Canadian Assessment of PL (CAPL-2), Pupils Health Knowledge Assessment (PH-KA); Scales to measure knowledge and preference of diet and PA (SCALES)

4.5 Discussion

This chapter demonstrates that there are currently assessment tools available in the literature that are both feasible *and* psychometrically appropriate to assess the sub-elements of the physical, affective and cognitive domains within the PL concept. Additionally, it is a finding from this review that there are tools that are appropriate for use within children aged 7-11 years that can also be administered within a primary school setting. However, no assessment tool captured the PL concept in its entirety, this may be attributed to the complex nature of PL and the ongoing debate concerning the underpinning philosophy, nature and definition of the concept (Whitehead, 2019; Durden-Myers, Whitehead and Pot, 2018; Lundvall, 2015; Shearer et al., 2018).

Study characteristics: The compulsory education sector, specifically PE, has become an advocate for PL due to the links with high quality PE (Edwards et al., 2018; Green, Roberts, Sheehan and Keegan, 2018). It is therefore encouraging that a total of 49 (physical, 19; affective, 25; cognitive 5) assessment tools have been conducted within a school setting. This continues to support the notion, that PL can be assessed and fostered within the primary school setting (Tremblay and Longmuir, 2017; Cairney et al., 2019). In addition, the findings from this review report that both cognitive and affective tools were typically conducted within a classroom setting i.e. paper-based questionnaires. While physical assessments tended to utilise an indoor gym hall or sports facility, likely due to the practical nature of each assessment within the physical domain. However, one tool, the Response to Challenge Scale, despite being a cognitive assessment measuring self-regulation, was administered as a practical assessment. The assessment is circuit based and leads children through a range of physical tasks which become increasingly difficult as they progress (Lakes and Holt, 2012). The aim of the assessment is not necessarily for the children to complete the physical tasks competently, instead their ability to self-regulate and overcome challenge is assessed. While

Becker et al. (2014) describes a positive relationship between self-regulation and participation in PA, it is perhaps the opportunity to be active which sets the RCS apart from other assessments within the affective and cognitive domains. Specifically, within the affective and cognitive domain 84% of assessments encouraged children to engage in sedentary behaviour through completing surveys, questionnaires and participating in interviews. However, the sole purpose of the PL concept is to encourage participation in regular PA, therefore utilising an assessment that provides opportunities for children to be active, is consistent with the nature of the PL concept. Moreover, the RCS is strongly aligned to the underpinning philosophy of monism which encourages the mind and body to be established as equal partners working in unison (Whitehead, 2019).

Another assessment which has embraced an alternative setting is the PA Research and Assessment tool for Gardening Observation (PARAGON) (Myers and Wells, 2015) which utilises an outdoor garden setting within a school. This provides an innovative example of a tool within the physical domain being administered in a different environment. Notably, the PARAGON assessed movement patterns that were specific to this activity and location, i.e. watering plants/digging/gripping/bending and stretching (Myers and Wells, 2015). By embracing a PA rather than sport specific context this provides an alternative to the typical fundamental movement assessment. This is a timely approach for/of what, as Whitehead advocates that varying environments play a pivotal role in the physical domain, providing challenging situations and opportunities for individuals to excel within their PL journey (Whitehead, 2010; 2019). A key recommendation of this review is that varied environments can be explored and utilised where possible for conducting PL assessments. This may allow assessors to consider an individual's potential to be active regardless of setting- continuing the idea of participation throughout the life course (Lundvall, 2015; Shearer et al., 2018; Green, Roberts, Sheehan and Keegan, 2018).

PL elements: It has been long proposed by Whitehead that an individual's PL journey is deeply personal and unique, therefore it may not be necessary to consider the concept as a whole, instead assessing the elements that link directly to health promotion i.e. fitness (Dudley et al., 2017; Cairney et al., 2019). Each tool included within the review assessed at least one PL element and as such there are 59 existing tools within the literature than can be utilised to support PL assessment within the 7-11 age range. Within the results of the physical domain all tools provided assessments for land-based movement skills; none of the tools assessed water-based activities despite swimming being the only compulsory PA within UK, Australian and American primary PE curriculums (Department of Education, 2017; Lynch, 2015). Additionally, many of the sub-elements that Whitehead describes as physical competence are already prominent within the PA and motor competence research fields. This was further supported by the findings of this review as a high proportion of existing tools assessed a range of fundamental movement skills and fitness components. Despite literature within the PL field warning against focusing too heavily on the physical domain (Whitehead, 2019; Robinson and Randall, 2018) it should still be a key consideration within assessments of PL. Specifically, research within the motor competence field reports that higher functional movement skill scores from assessments are directly linked with PA levels in children (Logan et al., 2015). Recent research supports this further by suggesting that children with higher motor competence scores are more likely to achieve PA guidelines (DeMeester et al., 2018). Consequently, while assessments should not focus solely on the physical domain it is an essential aspect of PL that should be assessed alongside the cognitive and affective domains. Moreover, the findings of this review also report a significant lack of tools available to assess the elements of aesthetic/expressive movement, sequencing, progression and application of movement specific to the environment, and this should be considered when making judgements within the physical domain. With regards to future tool

development, the missing elements within the physical domain should be considered in order to capture a complete picture of children's physical competence.

The findings related to the affective domain report that the element of motivation most frequently assessed, with 36% of assessments including a measure of motivation. This is a particularly positive result as alongside being crucial to the concept of PL, research highlights that motivation can underpin enjoyment in PE, suggesting that meaningful assessment can encourage intrinsic motivation (Ruben et al., 2019). Furthermore, Whitehead (2010) describes an individual who embodies PL as exhibiting a 'joy of movement' (pg. 30, 2010). In addition, the elements of confidence, perceived competence and self-perception were also included within existing assessments with 27% of tools including an aspect of each. It is well established within the literature that children who report higher levels of perceived-competence demonstrate higher levels of actual motor competence, particularly in relation to object control within the 7-11 age range (Robinson et al., 2015; Brian et al., 2018). As such, being able to accurately assess children's perceived motor competence may also provide support to predict actual motor competence within primary school children, highlighting children who may need additional support. While the majority of the elements within the affective domain were represented within assessment, this review found no tool that included an assessment of empathy. The ACS (2017) outlined empathy as a key factor for individuals managing the psychological aspects of their PL journey, advocating that it was required during PA to help control and understand emotional responses (ACS, 2017). This may be of particular importance as within the 7-11 age group children are developing higher order processes such as empathy and sensitivity that relate directly to emotional intelligence. Whitehead (2019) considers empathy when describing the inclusive nature of PL, advocating that empathy should be shown to all, allowing all individuals to explore new physical experiences and challenges regardless of individual needs (pg. 68, 2019). In addition,

Whitehead (2019) details empathy and emotional sensitivity as attributes of PL highlighting that children will experience better relationships and feelings of belongingness if these are fostered within the curriculum. Notably, Bryant's Empathy Index (Bryant, 1982) is used within the psychology discipline alongside self- report measures or teacher ratings to measure empathy, within both adolescent, child and special populations (Lovett and Sheffiled, 2007) however, this was not captured within the outlined search parameter included within this programme of research.

All assessments within the cognitive domain, apart from the RCS (Lakes and Holt, 2012) included an assessment for knowledge and understanding of the benefits of PA. This is a novel finding from this review as knowledge and understanding of PA and health concepts is considered to not only be associated with improved PA behaviours (Cale and Harris, 2018) but also an essential component within Whiteheads interpretation of the cognitive domain (Whitehead, 2019). Notably, the CAPL-2 included seven of the 11 proposed sub-elements (Longmuir et al., 2015) within the cognitive checklist, potentially as this is the only tool within this domain to attempt to assess PL as a concept. In addition, the results of the PL alignment checklist within the cognitive domain indicate a lack of assessments relating to the sub-elements of creativity and imagination in application of movement and knowledge and understanding of tactics, rules and strategy. However, this is not surprising given the majority of these tools were not designed to assess PL. A key recommendation in the development of future PL assessments is to aim to incorporate these essential sub-elements that are not already representing within existing assessments.

Whitehead advocates the importance of considering the concept of PL as a complete entity (Whitehead, 2010; 2019). However, while it is important to align closely to this interpretation, PL is also in its infancy stages and we cannot predict how the concept will develop and change over time (Cairney et al., 2019). Subsequently, a key recommendation of

this review is to utilise the assessment tools currently available where possible and in terms of tool development, focus on the areas of PL which are not already represented within existing assessments.

Psychometric testing: A key inclusion criterion for this review was an aspect of psychometric testing, therefore each study has reported on *at least* one psychometric property. A key finding across all studies within the physical, cognitive and affective domains is that studies tended to either test or report on one or two psychometric properties (reliability and/or validity) in detail but did not address all aspects of COSMIN risk of bias checklist. The psychometric property most frequently assessed across all domains was reliability. For example, all studies within the physical domain reported psychometric testing directly related to reliability and typically in the required level of detail (91%), similarly, within the cognitive domain five out of six assessments reported on reliability. In addition, within the affective domain 88% of assessments scored 'very good' for internal consistency, an aspect of reliability (Robertson et al., 2017). These findings suggest that the wider PA research field may prioritise the reliability of an assessment tool above the other psychometric properties or that it is not common practice to report the full range of psychometric testing conducted within the published literature.

Recent guidance from COSMIN outlines that PROM development and content validity are the most important measurement properties to be considered when developing a new tool (Mokkink et al., 2018). Within the affective domain, while 97% of tools reported on PROM development, 57% of these received an inadequate score, with a further 34% receiving a score of 'doubtful'. In addition, within the physical domain, 16 tools reported aspects of PROM development and/or content validity, however only 6 of these were scored at an adequate level. However, within the cognitive domain most assessments failed to provide adequate detail on concept elicitation i.e. the methods used to identify relevant items

and/or how these items were piloted and refined and therefore, were marked as not reported for PROM development and content validity. This is particularly concerning as if assessments are not proven to be valid for use within the targeted population, then arguably, they are not appropriate (Welk, Corbin and Dale, 2000; Guyatt, Walter and Norman, 1987). COSMIN recommend that in order to achieve a 'very good' score the relevance, comprehensiveness and comprehensibility should be considered in detail for assessment tools. For example, ensuring that included assessment items are relevant and understood by the target population, are key stages of ascertaining sound content validity. This can be achieved by including participants in the developmental process and encouraging the sharing of experiences and opinions (Morely and Van Rossum, 2019). Subsequently, providing opportunities for co-creation while allowing the target population to guide the direction of assessment, this aligning with the existentialist and phenomological aspects of PL.

The remaining six psychometric properties of measurement error, structural validity, hypothesis testing, cross-cultural validity, criterion validity and responsiveness were typically not reported across all domains, physical (58%), affective (85%) and cognitive (92%). These findings may be largely due to the PA research field not reporting full range of psychometric testing within the published literature. In particular, a lack of responsive testing could also prove problematic as it may prevent changes over time being considered, this was found to be a similar cause for concern within a recent systematic review of the psychometric properties of gross motor assessment tools (Griffiths, Toovey, Morgan and Spittle, 2018). For the operationalisation of PL on a curricular and international level, responsiveness is an important psychometric property to consider and would be invaluable to charting an individual's PL progress over the course of their lifetime. It is suggested from the results of this review that the psychometric properties of physical assessment tools need to be reported in more detail, specifically the process of determining PROM development/content validity

should be tested and reported in a transparent manner, the results of which informing tool selection.

Feasibility: Considering the feasibility of each assessment tool in relation to a primary school context was an essential aspect of this review and this is the setting in which the majority of tools were conducted. Despite not commonly considered a traditional measurement property; feasibility can be viewed as one of the top factors to consider in assessment development (Robertson et al., 2017). As such, feasibility is of high importance when considering the appropriateness of an assessment, especially as teacher's note lack of time and training to be among many challenges of conducting assessments within the school context (Chróinín and Cosgrave, 2013).The findings of this review were generally positive in regards to the feasibility of conducting assessments across each domain within a school setting. Specifically, for assessments within the physical domain the time to complete, equipment required and space needed for the majority of assessments suggest that a primary school is an appropriate setting to complete physical competence assessments. Additionally, although the cognitive and affective domains typically did not report the time, space and equipment needed to conduct assessments, it is a fair assumption that these would be positive due to the majority of assessments being paper based questionnaires.

Another key finding from this review is that the majority of both physical and cognitive assessment tools would need to be conducted by a PE/sports specialist or a researcher with additional quantifications. However, within the affective domain this information was typically not reported. Therefore, an effective assessment would need to consider who would be conducting it to determine any potential training needed, ultimately, this would be an influential factor in the overall cost of the assessment. Edwards et al (2018) also highlighted the need to support teachers with continuous professional development in order to ensure that pedagogical processes regarding assessment, teaching and learning were

appropriate. While it may require additional resources to effectively prepare classroom teachers to administer assessments, enabling the teacher to conduct and interpret the results of a PL assessment is particularly important as a classroom teacher will relate to and understand their pupils on a deeper level than that of a researcher (Durden-Myers and Keegan, 2019; Mandigo, Francis, Lodewyk and Lopez, 2012). Therefore, judgements made concerning performance, progression and development during assessment will in theory, be well informed and connect in-depth with the existentialism and phenomenology of each individual child. This is an important consideration as equipping classroom teachers with an assessment that they cannot fully understand, score and subsequently interpret the results of can hinder the teacher's ability to provide useful feedback and support the child on their PL journey.

Another key recommendation of this review is that participant understanding is an area which warrants further investigation to enable effective implementation of assessment. Particularly within the physical domain participant understanding was poorly reported compared to that of the affective (37%, adequate or very good) and cognitive (50%, very good) domains. Specifically, it was found that only one tool within the physical domain, Playfun (Cairney et al., 2018) considered the participants understanding of the assessment tool. This suggests that there is room for improvement in including both participant and administrators in the tool development process, ensuring that the tool is relevant, easily understood and accepted by both teachers and children. Additionally, participant understanding may be a particularly pertinent factor for the effectiveness of a school-based assessment as it can ensure the assessment is relevant, easily understood and accepted by its users. This suggests that both participants and administrators should be involved in the development process.

Finally, in regards to feasibility this review considered each domain individually and scored the feasibility of assessments accordingly. However, Whitehead advocates the

importance of considering holistic assessment that represents the physical, affective and cognitive domain equally. Within this review CAPL-2 was the only assessment to assess an aspect of each domain and although not equally weighted, this tool provides the closest alignment to Whitehead's interpretation of PL. For example, while the individual assessments within the physical, affective and cognitive domains of CAPL-2 scored positively for time (very good ****) the reported combined time to complete all of the assessments is 30-40 minutes per child giving, classifying the assessment in the fair category (**). This is an important consideration as to align with Whitehead's articulation of PL a holistic approach is required, as such assessing each domain at one time point has the potential to affect the overall feasibility scores provided, requiring further investigation.

Strengths and Limitations

There are various strengths of this systematic review which include: (i) that it is the first review to consider and appraise assessment tools that are currently available within each PL domain, (ii) it is a detailed review that considers not only the alignment of the assessment to the PL concept but also the psychometric properties and feasibility of each tool. Finally, (iii) this review provides a systematic and transparent approach which will contribute to the limited empirical research within the PL field. The limitations of this systematic review include: (i) only papers published in the English language were considered, thus the assessment tools were primarily derived from the US, UK, Australia, Canada and Europe. Additionally, information relating to psychometric testing, feasibility and PL alignment may have been reported in another language, therefore not being considered for this review, (ii) To be included into the review, articles had to be published in a peer reviewed journal, therefore, tools developed by practitioners and used currently within schools may not have been represented in this review.

4.6 Conclusion

This study is the first to provide a systematic review of assessments available in the literature to specifically assess Whitehead's articulation of PL. Consequently, this work is essential for implementing PL, whilst also establishing links between similar constructs within PA research. Although PL is presented as a novel concept it is clear from the findings highlighted in this review that aspects of the physical, cognitive and affective domains are currently being assessed, using tools that have acceptable psychometric properties. In addition, while psychometric properties were generally not reported well within the literature, more attention, in particular needs to be given to PROM development and content validity within the field of PA research. Furthermore, while assessments are typically proven to be feasible within a school context, further empirical research is needed to consider the feasibility of the scoring and administration of assessment tools by teachers as opposed to research teams. The key recommendations resulting from this review process is to focus on developing or adapting assessment tools specifically for use within challenging environments that support the teacher or practitioner to be the assessor. Finally, this review provides information that can be used by researchers and practitioners to inform the selection or development of tools for the assessment of the physical, affective and cognitive domains of PL.

Chapter 5

Stakeholder perceptions of a PL assessment tool for children aged 7-11 years old

5.1 Thesis map

Study	Objectives
Study 1: Global interpretations of PL	• To collate, compare, and critically review existing definitions of PL from leading organisations implementing PL agendas around the world Key findings: • Seven international groups were identified as currently working to promote and develop PL, each operating with at least one identifiable definition of the concept • There is a lack of consensus and clarity between definitions and interpretation of PL, as such transparency is needed when reporting alignment to the varying PL concepts • Whitehead's definition (2001) of PL has been utilised or expanded upon within the differing interpretations, providing the rationale for anchoring the work included within this thesis to Whiteheads 'conceptualisation of PL
Study 2: A systematic review of assessments related to the physical domain of PL among children aged 7-11 years	• To systematically review the academic literature for tools to assess the physical, affective and cognitive domains of PL within children aged 7-11 years. • To explore and critically discuss each assessment tool within the physical, affective and cognitive domains to appraise its (a) psychometric properties; (b) feasibility for use within a primary school setting and (c) alignment to the PL concept. Key findings: • There are currently assessments available to assess elements within the physical, cognitive and affective domains of PL • Generally, the psychometric properties of tools were not well reported within the literature, more attention, in particular needs to be given to PROM development and content validity within the field of PA research. • PL assessments are typically proven to be feasible within a school context, however, further research is needed to consider the feasibility of the scoring and administration of assessment tools by teachers as opposed to research teams.
Study 3: Stakeholder perceptions of a PL assessment for children aged 7-11 years	• To explore key stakeholders (academics/practitioners, teachers, and children) views of current practice, future directions and effective implementation of PL assessment, through concurrent focus groups • To inform the development of a rigorous, authentic, and feasible PL assessment for use with children aged 7-11

5.2 Introduction

The work included within this thesis has so far focused predominantly on the availability of tools to assess the PL concept, typically in relation to researchers and organisation. However, more information is needed regarding what is currently happening in practice, with a focus on understanding practitioners, teachers and children's perspectives. As such, investigating the feasibility of an assessment tool is of vital importance when determining if the assessment is appropriate for use within an educational context (Barnett et al., 2019). Particularly within a primary school setting, specific contextual requirements should be explored in detail; otherwise tools are not used in practice. Despite the feasibility of an assessment being imperative to the successful implementation, it is often not reported or overlooked within the tool development stages (Klingberg, Schranz, Barnett, Booth and Ferrar, 2019). It has been proposed that there are eight general areas that could be considered by feasibility studies: (i) demand; (ii) acceptability; (iii) implementations; (iv) practicality; (v) adaptation; (vi) integration; (vii) expansion; (viii) limited-efficacy testing (Bowen et al., 2009). Based on recommendations from Bowen et al., (2009) research conducted into the feasibility of a physical literacy assessment should identify the most appropriate factors and methodology considering the setting, community, or population included within the study. Specifically, three described areas were deemed to be of particular importance in aligning with the primary school context; (a) acceptability (to what extent is a new assessment judged as suitable?), (b) demand (to what extent is a new assessment likely to be used?), (c) implementation (to what extent can an assessment be successfully delivered to intended participants?). Research grade tools assessing physical competence have generally not been accepted within a primary school context, due to being too lengthy, complex and requiring expertise to administer or feedback results that a typical classroom teacher or practitioner may not possess (Lander, Hanna, et al., 2017; Van Rossum, et al., 2019). Within PL, it has been suggested that the classroom teacher should be empowered to implement an assessment, due to the close

relationship formed between teacher-pupil, while also fulfilling required roles such as providing feedback and charting progress (Durden-Myers and Keegan, 2019). Enabling the teacher to administer and interpret the results of a PL assessment is particularly important as a classroom teacher will relate to and understand their pupils on a deeper level than that of a researcher (Durden-Myers, Whitehead and Pott, 2018; Mandigo, Francis, Lodewyk and Lopez, 2007). It is suggested that judgements made concerning performance, progression and development during assessment will in theory, be well informed and connect in-depth with the existentialism and phenomenology of each individual child. Therefore, understanding and learning from teacher perceptions of PL assessment is crucial as the teacher is likely to be administering the assessment.

Recent research suggests that an assessment for PL should be valid, reliable and trustworthy for a specific population (Barnett et al., 2019). While there are currently PL assessment tools available for use within a primary school population (i.e., Canadian Assessment of PL [CAPL-2] (Longmuir et al., 2018), there is a lack of information regarding the actual use of these tools in practice by practitioners. Key information regarding stakeholder's perceptions of the aforementioned tools (particularly non-specailised practitioners i.e. a primary classroom teacher) is currently unknown, therefore it is queried whether the tools are appropriate to support assessment of PL in practice. Furthermore, no information has been published to date regarding children's perceptions, understandings and experiences of PL assessments.

While it is often considered that expert opinion is crucial to informing the developmental process of an assessment tool, particularly at the design phase (Longmuir, et al., 2018; Barnett et al., 2019; Mokkink et al., 2018; Robertson et al., 2018) within a primary school context arguably the teachers, support staff and children possess the expertise. Therefore, it is imperative to consider not only experts but also children and teachers

perceptions of assessment. Through embracing principles of participatory research (sharing the power between researcher-participant and encouraging the discussion of problems and solutions (Cornwall and Jewkes, 1995; Gibbs et al., 2018), teachers, children and practitioners are empowered to play an active role in informing the development of an assessment tool that is relevant to their educational context (Baum, MacDougall and Smith, 2006). Additionally, for child groups the use of interactive tasks based on the write, draw, show and tell methodology (Noonan, Boddy, Fairclough and Knowles, 2016) encourages a child-centered approach which creates opportunity for the child perceptions to be communicated. These methodological approaches further ensure that the development phase is relevant to, guided by the target population and that all 'voices' are heard.

This study aimed to explore these key stakeholders' perspectives of PL assessments for use within a primary school context, with children age 7–11 years. Similar to previous studies in related areas (Morley, Van Rossum, Richardson & Foweather, 2018; Ni Chróinín & Cosgrave, 2013) the use of qualitative techniques allowed compliance with safeguarding policy (children) and construction of meaningful themes, which could inform the development of a PL assessment framework.

5.3 Methods
Research team and reflexivity

Due to the close nature of engagement with participants via the qualitative technique of conducting and analysing focus groups, personal bias is unavoidable (Braun and Clarke, 2019; Tong, Sainsbury and Craig, 2007). Therefore, the following information aims to provide the reader with the personal characteristics of the lead authors (CS and HG) that may have unintentionally guided the direction of this work. Two female PhD students (CS and HG) conducted and analysed the focus groups before sharing the findings via triangulation with the wider research team (LF, LB, ZK- personal characteristics in appendix). Both lead authors had approx. 2 years previous experience of conducting focus groups and utilising other qualitative techniques. Both HG and CS attended internally provided facilitator training to support the focus groups (3-hour session led by Dr Paula Watson- personal characteristics in appendix). Educational backgrounds included HG (BSc Sports Science, MSc Sports Psychology) and CS (BSc Sports Science). The two lead researchers had no previous relationship with the child or teacher participants, and similarly the participants had no prior knowledge of the researchers or their backgrounds. However, the expert groups (n=2) conducted at the IPLA conference both HG and CS are active members of the IPLA and had presented previous work at the conference therefore, a previous relationship had been established with the participants.

Participants

 A purposeful approach to sampling was taken with children, teachers and practitioners who have a strong stake in assessment within a primary school context with the aim of providing rich, relevant and diverse responses (Giacomini and Cook, 2000; Tony, Sainsbury and Craig, 2007). Adult participants were informed that their involvement would be anonymous throughout the study and signed informed consent was obtained from each participant prior to commencement. For child participants gatekeeper consent, parental consent, and child assent

was also collected. Upon accepting the invitation, potential participants were asked to complete a demographic information form (gender, age, length of experience in teaching or current role, their role in school, knowledge of PL). The research was granted ethical approval by the Research Ethics Committee of Liverpool John Moores University (Ref. 18/SPS/037).

PL Experts

Physical literacy 'experts' included practitioners (N=8) and academics (N=13) who identified as having an interest in physical literacy assessment. Two focus groups were recruited as a convenience sample at the 2018 IPLA conference, whilst the third focus group was recruited via publicly available email addresses of academics/practitioners known to work within physical literacy and subsequently conducted at Oxford Brookes University. Participants were informed that their involvement would be anonymous throughout the study and signed informed consent was obtained from each participant prior to commencement. In total approx. 40 participant's experts were invited to take part in the focus groups. Twenty-two experts agreed to take part with one participant not attending the session due to travel delays. Reasons for refusal to take part were mainly due to time constraints and lack of availability within our data collection period. A total of 21 participants were included in the experts focus groups (female, n=14; male, n=7) and ages ranged from 25-60 years. Further, all participants classified themselves as working within education, sport or research, with a minimum of three years' experience of working within that field. Two participants self-identified their PL experience level as 'expert', six as 'proficient', nine considered with themselves as 'competent' and four further participants identified as a 'beginner'. For each session only the participants and researcher (CS or HG) were present during the focus groups. Participants typically did not have a prior relationship; therefore a short ice breaker task was utilised to give participants an opportunity to introduce themselves and make themselves comfortable

within the group setting (Krueger and Casey, 2002). All focus groups were conducted in English.

Teachers

Schools were contacted via publicly available email addresses and through school lists with agreed permission for contact by the University with regards to research studies. Information packs, explaining the purpose of the study were addressed to the school head teacher containing a request to forward to any staff member who regularly delivers PE to take part in focus groups. Initial emails included a participant information sheet providing details of the study and a gatekeeper consent form for the head teacher to complete and return. Head teachers were asked to respond directly to the lead researcher (CS) via email or telephone. The lead researcher (CS) made follow-up telephone calls to each school if a response was not received from the initial invitation. The participant information sheet noted that participating members of staff would receive a £20 shopping voucher for taking part in focus groups (provided by Liverpool John Moores University Physical Activity Exchange, PhD candidate's annual project costs). Participant written informed consent and brief demographic information was obtained by the researcher prior to commencement of each focus group. In total 15 schools were invited to take part in the study with 13 accepting the invitation. Three schools initially refused due to lack of time and teacher availability, while a further 2 schools dropped out due to failing to agree a mutually convenient time slot within the data collection period. Within the 8 participating schools, 135 teachers and 115 teaching assistants (TA) were invited, with 24 agreeing to take part. A total of 16 teachers and 8 teaching assistants (female, n = 20; male, n = 4), agreed to take part in the study, but due to medical reasons, one of these teachers became unavailable. The focus groups were conducted at ten different primary school sites across the United Kingdom; Scotland (Glasgow) (n=2), England

(Warrington, Manchester and Shrewsbury) (n=6) and Wales (Llandudno and Cardiff) (n=2).

Each school varied in size, pupil demographic and socio-economic status. A complete

breakdown of school demographics can be found in Table 1.19.

Table 1.19 Breakdown of participating schools' demographic information

	School 1	School 2	School 3	School 4	School 5	School 6	School 7	School 8
Area, Country	Central Scotland	Central Scotland	West Midlands England	North West England	West Midlands England	North West England	West Midlands England	North Wales
School Type	Publicly Funded	Publicly Funded	Academy	Voluntary Aided	Academy	Community School	Foundation	Welsh Establishment
School Inspection Grade (OFSTED/Estyn/ Education Scotland)	Very Good	Very Good	Good	Good	Good	Requires improvement	Outstanding	2
Total Number of Pupils	362	836	121	232	325	401	288	248
Gender	45% Female	55% Female	49.6% Female	49.6% Female	47.7% Female	45.9% Female	50.3% Female	NR
SEN	35%	15%	3.3%	0%	2.2%	0.7%	0.3%	NR
Pupils with English not as first language	5%	5%	0%	2.5%	2.7%	38%	3.1%	NR
% eligible for free school meals	5%	35%	8.1%	43.1%	19.1%	43.9%	13.5%	5.9% (average in Wales 18.9%)

Only the participants and researcher (CS or HG) were present during the focus groups, additionally, participants had previous relationships due to working within the same school. Efforts were made where possible to conduct focus groups in quiet, isolated space where teachers could talk freely (Flores, and Alonso, 1995; Krueger and Casey 2002) typically this was an empty classroom. Additionally, no senior members of staff were invited to attend the groups in order to avoid power differentials among participants, this aiming to ensure that all participants could feel comfortable talking to one another openly (Krueger and Casey 2002).

Children

Following gatekeeper consent (head teacher), information packs containing child and parent/carer information sheets and consent forms were distributed to all eligible (aged between 7-11 yrs.) child participants at recruited schools by school staff. The participant information sheet noted that participating children would be entered into a randomised prize draw to be in with a chance of winning £100 of shopping vouchers to be used on PE and sporting equipment for their school. For the purpose of this study, a maximum of six consenting children (out of 76 who had retuned completed consent forms) from each class were randomly selected via lottery method to take part in each focus group. A total of 57 children aged 7-11 years (n = 32; male, n = 25, Male=8$\pm$0.7yrs) participated in 10 focus groups. The focus groups were conducted in a quiet space outside of the classroom, typically in the school library, with one of the lead authors facilitating the session (CS and HG). An ice breaker task was utilised (appendix 3) to give each child the opportunity to state their name and draw themselves playing a favourite game, the purpose of this was to identify the children's voices on tape and also make the children feel comfortable and relaxed within the setting (Gibson, 2007). Additionally, no non-participants or members of staff were present. School premises were chosen for convenience, to provide a familiar location alongside allowing compliance with safeguarding requirements.

Data Collection

A semi-structured focus group guide was developed based on Bowen et al. (2009), specifically, three described areas were deemed to be of particular importance in aligning with the aims of the present study; (a) acceptability, (b) demand, (c) implementation. Whilst all questions aligned with the overarching research aims and the broad themes of demand, acceptability and implementation, the wording of questions were altered slightly for different participant groups (appendix 2). Informal field notes were taken (written and electronically) during each session and the researchers did not return the transcripts to participants for comment or correction due to time constraints. Focus groups with teachers lasted between 45-60 minutes to facilitate staff attendance either after school or during break time. Experts groups lasted approx. 60-90 minutes due to the time allotment during the IPLA conference and participant work schedules at Oxford Brookes. Focus groups with children were deliberately shorter in length and lasted between 30 and 45 minutes, this ensuring optimal participant engagement throughout (Gibson, 2007). All focus groups were audio-recorded using a digital Dictaphone.

For child groups the age appropriateness of included questions and related tasks were confirmed by researchers experienced in the 7-11-year age group (LF, ZK, LB) and approved by HCPC registered psychologist (ZK). Additionally, facilitator interview guides were piloted within two child focus groups to ensure that the duration and all language used was appropriate; the guide was further refined by reducing the length of the sessions (from 60 minutes to 30 minutes) and changing some complex wording of questions based upon recommendations by the children (appendix 1D). As the discussion continued an adaptation of the Write, Draw, Tell Method (Noonan, Boddy, Fairclough and Knowles, 2016) was utilised (appendix 1D) encouraging children to draw and discuss their experiences in P.E or a PA setting.

Figure 1.3 Example of a focus groups session

Within the adult (academics/practitioner and teacher) focus groups, questions followed similar themes. Firstly, discussing current experiences of assessment of PL and/or in PE, and positive and/or negative aspects. At this point, participants were given a sheet of paper or listing common barriers faced by when assessing in PE. Participants were asked to rank these barriers by perceived importance, and to talk aloud about their reasoning. The facilitator then prompted the group to explore potential ways to overcome these barriers and to discuss what an 'ideal assessment' would look like. Throughout the focus groups, participants were provided with pens and large sheets of paper to write down/draw any ideas and were prompted to use these at various points if they felt the need to.

All focus groups were audio-recorded using a digital Dictaphone and transcribed verbatim. With the child focus groups, autonomous engagement was encouraged by offering children choices (e.g. choosing where to sit) and nurturing a supportive relationship providing opportunities where children could voice their needs and opinions (Domville, Watson, Richardson and Graves, 2018). Children were positioned around the moderator in a circular position to project a non-authoritarian climate (Gibson, 2007) and it was clarified the moderator was not a teacher, reinforcing there were no incorrect answers and the children

could talk freely. The moderator at times directed questions to participants to involve quieter group members and ensure that all participants had the opportunity to express themselves.

Data Analysis

Focus group data (258 pages, Ariel size 12, double spaced; children pages=87, teachers pages=103 and experts pages=68) was transcribed verbatim and each transcript read several times by the two lead authors (CS and HG) and subsequently uploaded within NVivo analysis software (version 12, QSR International) for data handling. Participants did not have access to the transcripts therefore did not provide any feedback on transcripts. Thematic analysis was chosen as the preferred analytical technique as the analysis process allows the researcher to identify, analyse and report patterns (themes) within a particular data set (Braun and Clarke, 2019; 2006). Transcripts were initially analysed through a deductive process using Bowen et al. (2009) as a thematic framework which reflected the underlying research objectives and questions. An inductive process was also used, enabling additional themes to be represented that had emerged from discussions. This process subsequently required the reading of individual transcript in order to assign broad thematic codes; these then organised into higher and lower order themes. Verbatim quotations were also taken directly from the transcripts in order to expand upon these themes within the findings. This hybrid approach of inductive and deductive analysis provided a thorough exploration of the research questions by comparing existing stakeholder perceptions of assessment, as determined by the research questions as well as allowing for the development of new themes (Boeije, 2010).

A pen profile approach was adopted due to visual profiles being considered appropriate for representing analysis outcomes from large data sets via a diagram of themes (Ridgers et al., 2018; Aggio, Fairclough, Knowles and Graves., 2016; Mackintosh et al., 2011). Self-defining, verbatim quotations and frequency data were also used to expand the pen profiles. Participants were anonymised using pseudonyms. For profile inclusion, the

threshold was set at a minimum of 25% in consensus of a particular theme (Ridgers et al., 2018), with themes not reaching consensus reported within the narrative. In the present paper, data from different stakeholder groups is presented separately.

Methodological rigour

Recommendations made by Smith and McGannon (2018) regarding qualitative methodology guided data collection and analysis. Specifically, as theory free knowledge is not possible, the second author (HG), acted as a critical friend (Smith & McGannon, 2018). The second author (HG) independently reverse tracked the data analysis process from pen profiles to themes, codes and transcripts, and any discrepancies (n=37) were noted and discussed until a richer reading of the data was reached (CS and HG). The pen profiles and verbatim quotations were then presented by the first author to the research group, by further means of cooperative triangulation (Mackintosh et al., 2011). The authors critically questioned the analysis and cross-examined the data providing opportunity to explore, challenge and extend interpretations within the data (Ni Chróinín and Cosgrave, 2013). Methodological rigour, credibility and transferability was achieved via verbatim transcription of data and triangular consensus procedures. Dependability was demonstrated through the comparison of pen profiles with verbatim data and triangular consensus processes.

5.4 Results

Within each stakeholder group the higher order themes of implementation, acceptability and

demand were present but how these manifested differentiated between groups. Pen profiles

representing the different perceptions are displayed in **figure 1.4** (children), **figure 1.5**

(teachers) and **figure 1.6** (practitioners). Themes are represented within pen profiles if

consensus is reached (more than 25%). Within each of the key themes, participants'

experiences and perceptions are discussed within the emergent sub-themes.

Figure 1.4 Pen Profile representing children's perceptions of PL assessment
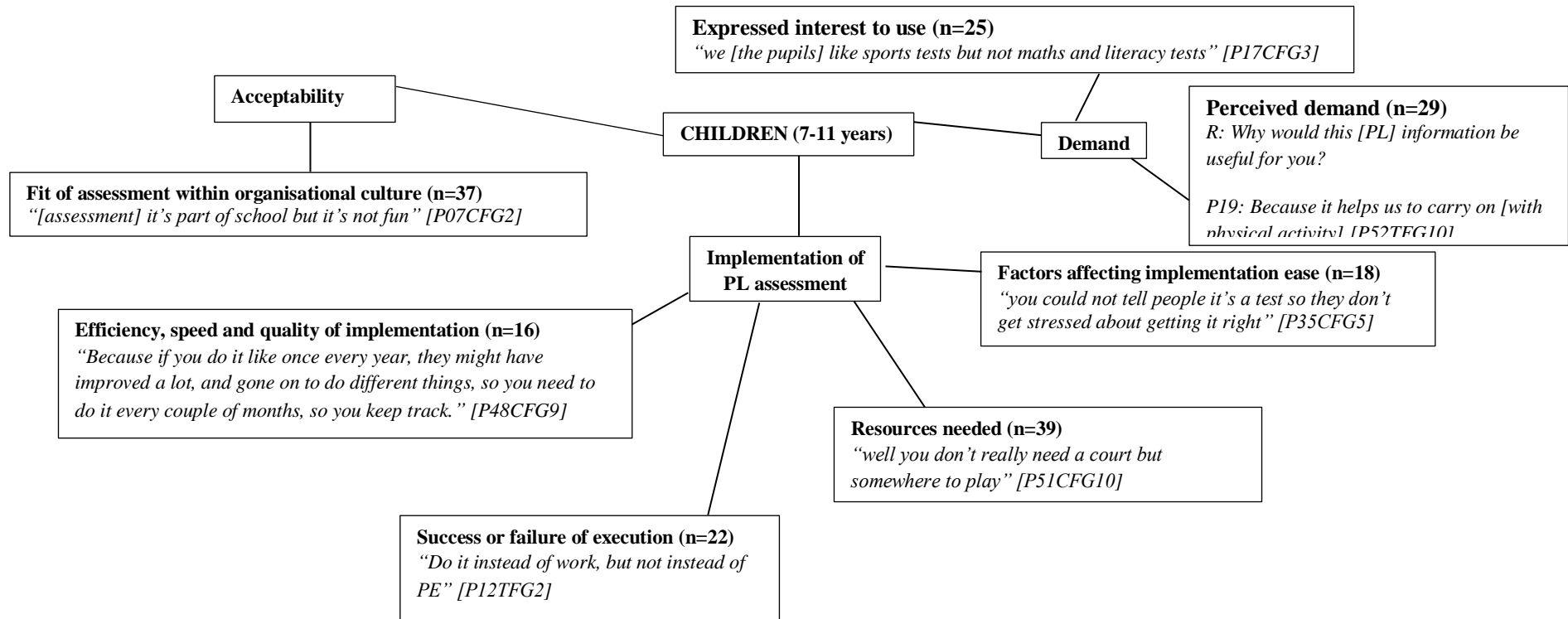

Expressed interest to use (n=25)
"we [the pupils] like sports tests but not maths and literacy tests" [P17CFG3]

Acceptability

CHILDREN (7-11 years)

Demand

Perceived demand (n=29)
R: Why would this [PL] information be useful for you?

P19: Because it helps us to carry on [with physical activity] [P52TFG10]

Fit of assessment within organisational culture (n=37)
"[assessment] it's part of school but it's not fun" [P07CFG2]

Implementation of PL assessment

Factors affecting implementation ease (n=18)
"you could not tell people it's a test so they don't get stressed about getting it right" [P35CFG5]

Efficiency, speed and quality of implementation (n=16)
"Because if you do it like once every year, they might have improved a lot, and gone on to do different things, so you need to do it every couple of months, so you keep track." [P48CFG9]

Resources needed (n=39)
"well you don't really need a court but somewhere to play" [P51CFG10]

Success or failure of execution (n=22)
"Do it instead of work, but not instead of PE" [P12TFG2]

Figure 1.5 Pen Profile representing teachers' perceptions of PL assessment

Intent for continue use (n=5)
"The challenge is to get kids to understand why they do it, so they become valuable and used to it" [P11TFG3]

Expressed interest to use (n=8)
"Making sure you give [assessment] the time it deserves" [P22TFG6]

Perceived appropriateness (n=6)
"Part of the problem is that adults underestimate what a kid knows or tries to introduce concepts that they know" [P02TFG1]

Intentions for continued use (n=5)
"[Assessment] supports the evidence… because then we've got evidence to back up, and we've got it written down" [P19TFG5]

Acceptability

TEACHERS

Demand

Perceived demand (n=9)
Progression of skills, if we're not sure where they're at or where they need to get to" [P07EFG2]

Fit of assessment within organisational culture (n=12)
"Even parent's night, the report cards, in the report section there's a huge section for Lit and numeracy but health and wellbeing are neglected" [P16TFG5]

Factors affecting implementation ease (n=7)
"There's no way I would feel confident assessing any [fundamental movement skills]" [P06TFG2]

Efficiency, speed and quality of implementation (n=14)
"I don't think it should be an external person, because I don't think they know the children well enough" [P13TFG3]

Implementation of PL assessment

Success or failure of execution (n=6)
"If it's on paperwork, you haven't got a chance" [P20TFG5]

Resources needed (n=8)
"QR code or something that the teacher can watch, so that they know what they're looking for" [P13TFG4]

Figure 1.6 Pen Profile representing experts' perceptions of PL assessment

Deductive themes

Acceptability

Within the child focus groups, participants accepted that assessment is a part of their school life, although at times they expressed concerns with the assessment experience.

"it's part of school but it's not fun[P19CFG4]"

Children valued an assessment that provides a fun experience but were particularly apprehensive about formal assessment situations where they may be judged by peers or asked to complete tasks individually in front of others:

"You think people are going to laugh at you or make fun of you when you're doing it [the assessment] by yourself" [P13CFG3]

Teachers perceived the acceptability of an assessment being linked closely to approval or support from management staff within the school. It was expressed that if the senior management team does not embrace or value PL, then it will be challenging to conduct any assessments, regardless of its perceived value, nature or feasibility:

"If your head teacher hasn't got that mentality, then it's doomed, I think." [P1GFG01]

Teacher groups also commented on the challenges associated with conducting assessments in school and reported that existing assessments were not able to capture the holistic nature of PL across the children's time at school.

"But most of our assessment tools - as you pointed out - are trying to show improvement in kids in schools, which is hard…but if it's a long term thing, if we could get some sort of charting progress tool or whatever that kids can engage with" [P18GFG02]

Similarly, from an expert perspective, it was reported that assessments should embrace novel techniques that provide the child with an impactful experience alongside collecting data to support them on their PL journey.

"I think it could be seen as an assessment for learning… in a non-conventional or creative way" [P17EFG3]

Demand

The results of the focus groups highlight that all stakeholder groups expressed a perceived lack of assessments within the PL domains across all ages within primary schools. Specifically, children commented that they have never been knowingly assessed in a PL context. However, when asked on the perceived demand of a hypothetical assessment tool, children vocalised that it could support the teacher in selecting an appropriate level of challenge and that this was needed.

> "[An assessment could…] let the teacher know what I am good at and what I am not, don't want things to be too easy or hard…then if she [the class teacher] had all this information, she would know what we're capable of, like not push us to [do] what we can't do." [P47CFG9]

Additionally, while the teacher groups also expressed that there is a current lack of assessments, they also commented on the lack of guidance surrounding the PL concept. Subsequently, teachers were not confident or comfortable with assessments related to physical education and PL as there was no support currently in place to guide this process. Lack of teacher education training was suggested as a possible reason for this.

> "I think that's because teachers generally, when they're at university, they just do [physical education training]...is it six hours of something? That's all... a little bit, but not a lot." [P20TFG6].

Due to the lack of training and ability for teachers to develop within their own PL journey, teachers expressed that they were confused and would not be able to confidently asses PL within their class.

> "we're not sure where they're at or where they need to get to" [P07EFG2]

Practitioner groups also expressed a clear demand for assessment in order to chart progress in children and inform best practice within the education sector.

"Do we [experts] think it's important to assess PL?

P1: Fundamental

P2: Absolutely fundamental [EEG1]"

Implementation

Throughout the participant groups, it was communicated that stakeholders were not satisfied with existing assessments. It was expressed across all groups that PL assessment rarely happens within primary schools and if it does, assessment focuses solely on the physical domain.

"I don't think it's enough to run to the green hoop then that's the end of the story. Why are you standing there, what skills did you use? That's a simple way to assess it but you have to follow up after" [P15TFG3]

With regards to implementation both children and teachers suggested the use of self-assessment but for different reasons; teachers to reduce assessment burden and children to allow them to chart progression overt time. In particular, it was highlighted within child groups that self-assessment often already happens informally out of school hours, with children using technology such as IPADS or iPhones to record the process of learning a new skill.

"Record yourself doing that [new skill] every day, and then as you start to do it, you'll get better every day" [P36CFG7]

Additionally, teachers advocated the use of self-assessment, peer assessment and the use of technology to reduce the assessment burden on teachers or support staff. It was also suggested that through self-assessment and peer assessment children can reflect and support

one another, this prompting children to take ownership of their relationship with physical activity:

"What about video evidence, so some children videoing each other and watching it back and assessing it together. Teachers don't need to do it, the kids can. They can talk together, making it less formal. Motivating each other" [P02TFG1]

Similarly, within the practitioner groups self-assessment was suggested and the importance of reflective practice to guide self-assessment. Using longitudinal assessment to capture changes and provide support over time was also expressed as being important while ensuring that constructive and meaningful feedback was translated to the children.

"it's your balance of time to assess, and talk… and to reflect and to teach and to move on" [P09EFG1]

Inductive themes

Within the child focus groups, inductive themes emerged from the discussion, mainly due to the flexible nature of the semi-structured groups providing the opportunity for children to lead the direction of the conversation (Braun and Clarke, 2012). The inductive themes of; role of others, self-awareness and feedback emerged within this age group. The pivotal role that others play in the assessment process was discussed frequently within the child groups, with children suggesting that family, friends, teachers, coaches and they themselves could support and administer assessments. However, it also emerged that if another individual was conducting or supporting the assessment, they had to be trusted to provide a positive and safe experience, the children generally expressed this is why they didn't like previous assessments.

"so teachers just make it like really tense sometimes…it's like this is your grade 2 assessment, if you don't do it you'll have to go back and revise all this stuff again" [P51CFG10]

Children of this age group were extremely self-conscious and therefore, wary of being humiliated or embarrassed in front of others during an assessment. Additionally how children give and receive feedback was also discussed on numerous occasions, with a consensus being reached that only constructive feedback should be given. Children articulated the impacts of receiving negative feedback that could have the potential to make them feel "sad" or "disappointed" in themselves. Even in situations of failure it was implied, children advocated that a supportive, nurturing environment was always key to enjoying the assessment experience.

> "well if the teacher was kind they would say well done, very good you tried your best…next time try and practice more" [P28CFG5]

Challenges

During the focus groups there were times where consensus was not reached and/or participants found questions/tasks challenging. In particular, the WDST task described earlier within this chapter, while evoking powerful discussion, the participants found it difficult when asked to provide specific details regarding an "ideal hypothetical assessment". Additionally, within the child groups, participants found questions regarding the cognitive and affective domain challenging.

5.5 Discussion

The aim of this study was to explore stakeholders' views of current practice, future directions and effective implementation of PL assessment within primary schools. Bowen et al (2009) suggests that there are eight general areas that could be considered by feasibility studies: (i) demand; (ii) acceptability; (iii) implementations; (iv) practicality; (v) adaptation; (vi) integration; (vii) expansion; (viii) limited-efficacy testing (Bowen et al., 2009). Within a physical literacy context, research conducted into the feasibility of a physical literacy assessment should identify the most appropriate factors and methodology considering the setting, community, or population under study. The findings from this study will be used to guide the development of a rigorous, authentic, and feasible PL assessment tool for use within the 7-11 age range. To the best our knowledge, this is the first study to use qualitative techniques to investigate stakeholders' perceptions of PL assessment. Notably this is the first study to include children as stakeholders. All stakeholder groups viewed the assessment of PL as important but could only identify assessments within the physical domain. This suggests a significant gap in the assessment of the affective and cognitive domains within primary schools. Additionally, all stakeholder groups highlighted the importance of embracing technology and self-assessment/reflection as part of an assessment process. This was posed as a potential solution to reduce the time spent on administering assessment while giving children an opportunity to take a leading role in their own assessment of their PL journey. The following discussion is sectioned into the higher order themes identified, triangulating perspectives across the three stakeholder groups.

Acceptability

Within this theme, both teacher and expert participant responses indicated that the themes of *fit within organisational culture* and *perceived appropriateness* were of importance. While children and experts reached consensus for *satisfaction*.

It was clearly expressed across all stakeholder groups and within the supporting literature that there is a demand for a PL assessment to be available for use within primary schools (Barnett et al., 2019; Edwards et al., 2018; Giblin, 2014). While assessment of PL does not appear to be common practice within the primary school setting, at present, during the focus groups children could give examples of assessment of the physical domain.

"we've been assessed like physically, to know what we can do and what we can't do" [P70CFG4].

This echoes literature that suggests that the physical domain is at times the only PL element assessed and considered within existing assessments (Edwards et al.,2018). Despite this, the findings from the focus groups highlight that both teacher and practitioner groups value the holistic nature of the concept and the importance of considering all PL elements, this supporting Whitehead's interpretation of the concept (Whitehead, 2010; Durden-Myers, Whitehead and Pot, 2018). In particular, teachers reported the importance of considering the cognitive and affective domains as well as the physical.

"It might be important to assess the physical stuff but also the other stuff as well, the knowledge or the motivation" [P15TFG4].

Both teacher and participant groups advocated that the term PL is being used more frequently within an educational context and that an assessment would be accepted within the primary school context to support best practice but that it should consider all of the PL domains.

Within the child groups, it was clear that above all else, children valued the importance of a fun and enjoyable experience in regard to participating in assessments. An essential factor for promoting acceptability within the child groups was simply, would the experience be fun? In addition, recent investigation into experiences during PE found that for children having fun can be the sole motivation for participating in PE rather than health and well-being purposes (Walseth, Engebretsen and Elvebakk, 2018). Furthermore, it was also

reported that having fun is strongly linked to the meaningfulness of PE and that enjoying the

activities during PE gives them a sense of purpose (Chróinín, Fletcher and O'Sullivan, 2018).

While fun was reported as a main theme throughout all the child groups it is important to

consider that 'having fun' is a subjective experience for many and that it is challenging to

create a fun environment while also having a focus on learning (Walseth, Engebretsen and

Elvebakk, 2018). In regard to PL assessment investigating what assessment practices would

promote opportunities for fun is key, while recognising that similar to PL experiences what is

considered to be fun and enjoyable is unique and children should be encouraged to explore

this through individualised approaches (Green, Roberts, Sheehan and Keegan, 2018).

While it was clearly expressed across all groups that there was value to having a PL

assessment tool, the adult groups particularly the teachers commented on the challenges of

regarding acceptability within the primary school context. Teachers commented that if the

assessment did not receive support from senior management team then regardless of their

efforts, it would ultimately not be accepted within the specific school context. This mirrors

findings reported from research investigating school 'politics' and the hierarchical power

systems within schools, namely that engaging the head teacher was crucial (Schuck,

Aubusson, Burden and Brindley, 2018). Therefore, it is arguably, important to consider not

only the teacher who is administering the assessment but also the management team or head

teachers who would be influencing the use of said assessment. In the context of our study

head teachers had to provide consent for their school to take part in the study, potentially

valuing PL and related constructs however, this may not always be the case (Skinner, Leavey

and Rothi, 2018). Assessment within the primary school context is common practice (Hay

and Penny, 2012) and although this is typically seen as positive practice to inform learning

outcomes, expert groups described 'evaluation fatigue' within schools and stressed that

assessments must have a strong and relevant purpose in order to be accepted within a primary

school context. Previous examples of fruitless assessments were outlined along with the frustrations that accompanied these.

"Don't you guys get evaluation fatigue? I just get 'ugh not another evaluation form'.

It comes everywhere you go and it becomes a pain and I'm bored of it" [P1EFG1]

In order for PL experts to accept assessments, the tools should strive to provide support and context to an individual's PL journey while not adding to the already overburdened teacher role.

Demand

Within this higher order theme, both lower order themes of perceived demand and existing assessments reached consensus across all participant groups. Teacher and expert focus groups also reached consensus for expressed interest of intent to use.

There was a clear perceived demand for a PL assessment across all stakeholder groups. In teacher groups it was suggested that an assessment could be used to provide evidence and support best practice. In addition, the results of an assessment could compliment reporting structures and feedback useful information to the children, parents and other staff members. However, teachers also reported a significant lack of training related to administering assessments related to the PL domains. It was highlighted that while there was demand for an assessment within the teacher group's considerable support and training would be required in order to conduct assessments successfully. Similarly, practitioner groups also expressed that teachers would require additional support to assess and understand the concept on the required levels to support children on their PL journey; this potentially affecting the feasibility of an assessment tool.

"One thing I've noticed when I've worked in schools is teachers themselves don't

have very good PL. So it then makes it difficult for them to make a judgement on a

child and how their PL journey is…so is what you're going to get back even going to

be accurate?...teachers don't get training on a lot of the PL stuff which is an issue"

[P09EFG1]

Despite the perceived challenges surrounding enabling teachers to assess the concept it was suggested by both practitioner and teacher groups that a teacher or school staff member was a preferred choice for an assessor, rather than that of an external coach or researcher.

"they don't have that whole view of the child… They don't know like the family background, they don't know what Mum and Dad are doing, or anything that's going on at home" [P13TFG3]

It was discussed at length within practitioner groups that teachers due to the nature of their role had a greater understanding of children's individual context's and interactions with the environment. While it may be a challenge for teachers to learn, understand and ultimately assess the physical concept, it is by no means an impossible feat and it could be an alternative opportunity to align, in parts with the underpinning philosophies of the concept. Furthermore, the findings from this study support recent research that highlights the considerable need to support teachers and practitioners to further understand and implement PL practices (Edwards et al., 2019). It is suggested that continuous professional development and ongoing training can be utilised to support the PL in schools, as this type of professional development is common practice within schools (Lander et al., 2017). Additionally, recent research from Edwards et al. (2019) reports on the success of utilising professional development programmes and PL workshops to increase teachers' knowledge of, and operationalisation of, PL principles. It was reported that attending regular professional development sessions successfully improved teacher understanding, knowledge and confidence in teaching PL (Edwards, 2019). It is essential the any PL training or continuous professional development sessions are relevant and engaging, inspiring teachers to connect with the concept further (Durden-Myers and Keegan, 2019). Lander et al., (2017) recommends that teacher training

sessions should be 'sustained' (i.e. one day or more), embrace a multidimensional approach while providing detailed content that includes pedagogy and translation into practice. Furthermore, it is crucial to that teacher satisfaction is measured within the training, the results of which guiding the development of future sessions (Lander et al., 2017).

Implementation

During the child groups, participants shared previous assessment experiences that they had taken part in while at school mainly in reference to key curricular subjects (i.e. maths and English) as assessments in physical education or in relation to PL were not common. Unfortunately, a significant majority of children reported negative experiences, where they felt under pressure and reported additional stress due to formal assessments and the consequences of not performing well.

"You're doing the test… it gets you like really worried and stuff" [P19CFG4]
This echoes a trend in the literature that suggests 'test anxiety' is becoming increasingly prevalent within the younger age categories due to high-stakes assessments being frequently administered (Putwain and Best, 2011).

It is also reported that a negative physical competence assessment experience, can have the potential to decrease confidence, self-esteem and motivation to participate in PA (Wiersma and Sherma, 2008; Welk and Blair 2008). Therefore, it is extremely important that a supportive and nurturing assessment environment is fostered as the sole purpose of PL is to engage and support children in PA participation for life (Whitehead, 2001; 2010; 2019). In particular, the potential for physical assessments to cause physical or emotional discomfort for children is discussed at length within the literature (Graser, Sampson, Pennington and Prusak, 2011; Wiersma and Sherman, 2008; Cale and Harris, 2009; Allender, Cowburn and Foster, 2006). In addition, a systematic review of qualitative research within the UK found that a negative experience during school PE was one of the strongest factors discouraging

future participation in PA (Allender, Cowburn and Foster, 2006). The results suggested that creating a safe environment where children can perform without fear of reprimand or embarrassment was considered to be a key motivation to participating in PA, alongside opportunities to experiment with and engage in usual activities (Allender, Cowburn and Foster, 2006). Furthermore, examples of how practitioners can achieve a positive testing environment include providing opportunities to practice skills before being assessed, performing skills as part of a small group and pairing children with similar abilities to work together (Cale and Harris, 2009). Additionally, children reported anxieties around performing in front of peers, particularly if the level of challenge was not perceived to be appropriate for their individual level.

> "So they [swimming teachers] know what lane to put you in, so they don't put you in the deep end if you're meant to be in the shallow end" [P03CFG1]

Within this age group, self-preservation is crucial, and children are becoming increasingly self-aware and influenced by peer's opinion, in fact this often influences the opinions they have of themselves (Amaya-Hernández et al., 2019). Despite the negative connotations of existing assessment, when asked directly about assessments in P.E or related to PL domains children perceived that they would enjoy these due to the active nature and previous positive experiences in physical education. Additionally, key components that are reported to influence children's enjoyment of PE are both peer and teacher behaviour during the session and tailoring tasks to the individual ability level of each child (Domville, Watson, Richardson & Graves, 2019). Subsequently, many of the students within the child groups associated taking part in P.E as a fun experience and welcomed the idea of an assessment in this context. Therefore, it may be not the assessment itself that children are wary of but the experience typically surrounding formal assessments where they are put under pressure to perform in public settings. Providing a positive, relaxed and informal environment while

administering assessments could be essential to ensuring children have a positive assessment experience, while enabling assessment for learning principles.

It has been well established within PL research that each individual's PL journey is personal and unique to them (Whitehead, 2019), and due to the underpinning philosophies of the concept cannot be easily charted or assessed by others (Green, Roberts, Sheenan and Keegan, 2018). Across all stakeholder group's self-assessment was suggested as a potential method of charting an individual's PL journey. Further, within the PA research discipline self-report questionnaires (or in the case of young children, proxy-report by a parent/carer) have been used previously to measure SB and PA levels in child populations (Atkin et al., 2012; Lubans et al., 2011). However, self- and proxy-report tools are known to be susceptible to recall errors, misrepresentations and social desirability bias (Atkin et al., 2012; Loprinzi and Cardinal, 2011). Therefore, there is a need to look at novel self-report measures to assess PL. Despite the potential challenges of utilising self-assessment methods within a primary school context it may provide a unique opportunity for children to reflect and understand where they are on their unique PL journey- allowing them to track their own progress. Across all stakeholder groups when self-assessment was suggested the importance of utilising technology was a considered as a means of facilitating this within a school setting.

As technology continues to advance, studies have demonstrated the effectiveness of utilising apps and mobile devices within assessment in PE (O'Loughlin, Chróinín and O'Grady, 2013; Penney et al., 2012; Morley and Van Rossum, 2019). Furthermore, technology has been reported to overcome many barriers to assessment such as time constraints, lack of space and engagement in teacher populations (Jones and Moreland, 2004; Morley and Van Rossum, 2019). Specifically, it has been reported that within the 7-11 age range, self-assessment within PE using video provided an engaging experience for students regardless of ability (O'Loughlin, Chróinín and O'Grady, 2013). Due to the authentic nature of the tasks, using

video to record and replay clips during PE provided students with a realistic and inspiring reflection of their performance (Penney et al., 2012).

Strengths and Limitations

To the best of our knowledge, this is the first qualitative paper to explore a diverse range of stakeholders' perceptions of PL assessment within a primary school context. A main strength of this research is the large number of participants included alongside the range of stakeholders perspectives collected, from children, teachers and educational practitioners. The application of focus groups across different stakeholder groups, within various geographical locations allowed for a rich understanding of how assessment is understood, valued and implemented in regards to the PL concept. The findings do not however allow conclusions to be drawn on specific assessments that are currently available; instead this is a general overlook of existing assessments and perceptions of assessment. Further research is needed to evaluate the feasibility of individual PL assessments and how successfully these can be administered with the results translated within a primary school context. While the authors felt data, saturation was reached (Guest, Bunce, & Johnson, 2006, p. 59) due to the extensive coding, back coding and triangulation process during thematic analysis, caution is warranted before generalising the study findings to all primary school settings. Specially, the findings do not allow for generalisation for different age groups, or assessments conducted outside of the school environment and this is a limitation of this work.

5.6 Conclusion

The findings within this study indicate that stakeholder groups recognise the demand for a PL assessment tool in order to support the concept and inform best practice. Additionally, it was recognised that in order to implement PL assessment within a school setting while enabling the teacher to be the assessor, considerable support will be needed. The assessment experience was considered in detail and stakeholder groups expressed the importance of providing a positive experience for children, where their relationship with PA can be

nurtured. Specifically, children valued the perceived enjoyment of assessments and recommended that these mirror positive experiences already occurring within the physical education sector. Across all stakeholder groups self-assessment was suggested as a suitable method of assessing PL, supported by technology. Further feasibility studies are needed to consider the implications of using self-assessment within the 7-11-year population and appropriateness of these methods in practice. This research gives an in-depth insight into stakeholder's perceptions of PL assessment and an opportunity for stakeholders to share their voices and experiences with researchers, this information will be reflected upon and utilised to guide the development of an assessment framework for use in both research and practice.

Chapter 6
Recommendations for the development of a PL assessment tool for use within children aged
7-11 years

6.1 Thesis map

Study	Objectives
Study 1: Global interpretations of PL	• To collate, compare, and critically review existing definitions of PL from leading organisations implementing PL agendas around the world Key findings: • Seven international groups were identified as currently working to promote and develop PL, each operating with at least one identifiable definition of the concept • There is a lack of consensus and clarity between definitions and interpretation of PL, as such transparency is needed when reporting alignment to the varying PL concepts • Whitehead's definition (2001) of PL has been utilised or expanded upon within the differing interpretations, providing the rationale for anchoring the work included within this thesis to Whiteheads 'conceptualisation of PL
Study 2: A systematic review of assessments related to the physical domain of PL among children aged 7-11 years	• To systematically review the academic literature for tools to assess the physical, affective and cognitive domains of PL within children aged 7-11 years. • To explore and critically discuss each assessment tool within the physical, affective and cognitive domains to appraise its (a) psychometric properties; (b) feasibility for use within a primary school setting and (c) alignment to the PL concept. Key findings: • There are currently assessments available to assess elements within the physical, cognitive and affective domains of PL • Generally, the psychometric properties of tools were not well reported within the literature, more attention, in particular needs to be given to PROM development and content validity within the field of PA research. • PL assessments are typically proven to be feasible within a school context; however, further research is needed to consider the feasibility of the scoring and administration of assessment tools by teachers as opposed to research teams.

Study 3: Stakeholder perceptions of a PL assessment for children aged 7-11 years	• To explore key stakeholders (academics/practitioners, teachers, and children) views of current practice, future directions and effective implementation of PL assessment, through concurrent focus groups • To inform the development of a rigorous, authentic, and feasible PL assessment for use with children aged 7-11 Key findings: • Stakeholder groups recognise the demand for a PL assessment tool in order to support the PL concept and inform best practice. • Children valued the perceived enjoyment of PL assessments and recommended that these could mirror positive experiences already occurring within the physical education provision. • Across all stakeholder groups self-assessment was suggested as a suitable method of assessing PL, supported by technology.
Chapter 6: Recommendations for the development of a PL assessment tool for use within children aged 7-11 years	• To outline findings from previous chapters to inform evidence-based recommendations for a PL assessment within the 7-11 age range • To outline assessment resources currently available, whilst highlighting potential areas for future PL tool development.

6.2 Introduction

Whitehead's recent work suggests that an individual's PL is a disposition that can be reflected on and described (Whitehead, 2019). Further, charting an individual's progress should be that of 'what is your current perception of participating in PA?' rather than 'are you physically literate?' (Whitehead, 2010, pg.72). As such, PL assessment is deemed by several authors to be crucial for supporting children within their PL journey (Cairney et al., 2019; Barnett et al., 2019). In order to chart progress on an individual and population level, appropriate measurement of PL is required with the results of such assessment informing best practice (Edwards et al., 2018). While the research conducted within this thesis focuses specifically on schools, due to the lifelong nature of PL effective assessment needs to be considered across the life course (Whitehead, 2019). Furthermore, assessment is a foci on which curricula are informed and standards measured, therefore to give PL longevity, purpose and credibility within a primary school setting, types of assessment must be considered - perhaps not only in a traditional (summative) sense but also by embracing assessment-for-learning (formative) approaches (Giblin, Collins and Button., 2014; Corbin, 2016). The primary focus of formative assessment is to educate and improve performance. This differs significantly from summative assessment whereby the goal is to measure student performance or learning (Black and William, 2003; Dixson and Worrell, 2016). Within this chapter both formative and summative assessment approaches will be discussed in detail in order to provide clear recommendations for a PL assessment for use within 7-11 year olds. In addition, this chapter will strive to outline assessment resources *currently* available within the literature, whilst also discussing potential requirements for future tool development. The purpose of providing pragmatic recommendations within this chapter is to encourage teachers, practitioners and children to work collaboratively to start assessing the PL concept within the primary school setting in order to develop a more comprehensive understanding of PL within the 7-11 child population.

6.3 Summative Assessment

The National Research Council defines summative assessments as "cumulative assessments that intend to capture what a student has learned, or the quality of the learning, and judge performance against standards" (National Research Council, 2007, p. 25). Due to the nature of summative assessments, they are typically administered at the end of a topic, unit or period of learning. Summative assessment is common practice within a primary school setting and as such teachers should be familiar with the assessment style. However, in terms of PL assessment it is crucial that the summative assessment experience should not only give pupils the chance to *demonstrate* their progress within their PL journey but also provide opportunities to solve problems or overcome barriers that they have encountered (Dixson and Worrell, 2016). As such, the following recommendations aim to encourage teachers, practitioners and children to work collaboratively to start assessing the PL concept within the primary school setting.

- **Recommendation 1: CAPL-2 should be used by teachers as a summative tool to assess children's PL**

The findings from study 2, collectively report that CAPL-2 (Longmuir et al., 2015) is deemed to be the best example of a holistic, user friendly and robust tool for use within the 7-11 age range. In comparison to the other assessments within this review, CAPL-2 displayed strong methodological rigour for many of the psychometric properties included within COSMIN. Specifically, CAMSA an aspect of the physical assessment scored as 'adequate' for PROM development and content validity. Furthermore, the relevance and comprehensiveness of CAMSA was rated as 'satisfactory' with the quality of evidence being deemed as 'high', this highlighting that the tool is fairly representative of the physical aspects listed within the assessment (Terwee et al., 2018). Similarly, the questionnaires provided to assess the affective and cognitive domains were also scored positively; affective (content validity and

structural validity rated as 'very good', cognitive (structural validity was also rated as 'very good'). When appraising the feasibility of using CAPL-2 within a primary school, across the affective and cognitive domains the questionnaires were scored as excellent (****) for both space and equipment, while CAMSA was rated as excellent (****) for both time and equipment. However, it is important to note that while the individual assessments scored positively, conducting the assessment as a whole may affect the feasibility scores i.e. the reported combined time to complete all of the assessments is 30-40 minutes per child classifying the assessment in the fair category (**).

In terms of PL alignment, CAPL-2 assesses the majority of the PL elements included within the bespoke alignment checklist (Table 4.1; Table). Specifically, the assessment provides tools for 50% of the proposed PL elements within the physical domain, 33% within the affective domain and for 7 of the 11 (64%) proposed sub-elements within the cognitive domain. However, it is also important to consider that CAPL-2 has been criticised within the literature for being heavily weighted towards the physical domain and this has been discussed at varying points within this thesis (Robinson and Randall, 2016). For example, CAPL-2 gives children the opportunity to score 30 points within the physical and affective domains and only 10 in the cognitive domains, thus affecting the overall PL 'score'. It is therefore suggested that when using CAPL-2, assessors strive to find opportunities to assess each domain *equally* (i.e. assessing each element within that domain), thus aligning with Whiteheads conceptualisation of the PL concept (Whitehead, 2019). It is important to consider that psychometric testing would be needed within any adaptations or new assessments. The results from the PL alignment work highlights that CAPL-2 does not represent each domain fully, i.e., within the cognitive domain the ability to identify/describe movement, creativity/imagination, decision making, tactics/rules/strategy and safety considerations are not being assessed. It is therefore, recommended that assessors should

strive to use existing tools alongside CAPL-2 to strive to capture a complete overview of children's PL. For example, the Response to Challenge Scale (Lakes and Holt, 2012) could be used alongside CAPL-2 to represent the missing element of decision-making. In addition, decision making is a very difficult element to assess and may need lots more work to examine the effectiveness of capturing this sub-element. Further, recent research suggests that various assessments can be utilised to measure the different elements of PL and that this should be encouraged where necessary (Barnett et al., 2019). However, within existing assessments CAPL-2 provides the closest alignment to Whitehead's conception of PL, demonstrates strong methodological rigour, has proven to be feasible for use within a primary school setting and is one of the few holistic assessments available. As such, CAPL-2 should be the starting point for PL assessment in children aged 7-11 years.

- **Recommendation 2: Providing a positive experience for children during summative assessments is fundamental to assessing PL**

A common criticism of summative assessment is that it typically creates 'high-stakes' scenarios where children are put under pressure to perform (Gardner, 2012). Additionally, within the child focus groups from chapter 5, when asked about summative or 'traditional' testing situations the children typically reported negative feelings of stress, unease or anxiety. Furthermore, the potential for physical assessments to cause physical or emotional discomfort for children is discussed at length within the literature (Graser, Sampson, Pennington and Prusak, 2011; Wiersma and Sherman, 2008; Cale and Harris, 2009; Allender, Cowburn and Foster, 2006). A systematic review of qualitative research within the UK found that creating a safe environment where children can perform without fear of reprimand or embarrassment was considered to be a key motivation to participating in PA (Allender, Cowburn and Foster, 2006). Examples of how practitioners can achieve a positive testing environment include

providing opportunities to practice skills before being assessed, performing skills as part of a small group and pairing children with similar abilities to work together (Cale and Harris, 2009). It is important for assessors to consider how existing assessments can be administered to achieve a positive testing environment, i.e., by setting up various obstacle course style stations teachers could encourage children to complete CAMSA in groups and give children an opportunity to practice the skills before the formal assessment potentially as part of a warm up exercise. The goal of any PL assessment should be to support continued engagement in PA and as such creating a positive assessment experience is crucial.

- **Recommendation 3: Future PL assessment development should focus on areas that are currently underrepresented in existing assessments**

An important finding from the work included within the systematic review is that some of the elements within the concept are not currently being assessed within the existing tools. In particular, the missing elements from each of the PL domains include: physical domain (movement skills in water, aesthetic/expressive movement, and progression), affective domain (empathy) and the cognitive domain (tactics, creativity/imagination and ability to identify/describe movement). However, Keegan et al. (2019) compares PL to that of the 'periodic table', defining both measurable and discrete elements, thus utilising a metaphor from the way that chemical elements can combine to form more complex compounds and mixtures. It is suggested that, like the periodic table, the elements are not definitive and as such, furthermore, elements of the PL concept may emerge over time. Therefore, in order to support the assessment of the concept it may be necessary to outline the key indicators of PL and focus assessment efforts around them. For example, a recent longitudinal study reported that for children aged between 6-9 years a high level of cardiorespiratory fitness was a determinant of improved coordinated scores (Santos et al., 2018). This could suggest that

there may be key assessments that relate to several elements of PL i.e. physical fitness could be assessed and this may lead to further information regarding other PL elements. Furthermore, this type of grouping or clustering is common practice within the PA research discipline where cluster/component analysis is used to identify hierarchical structures (Jago et al., 2018; Dumuid et al., 2018). In addition, within the CAPL-2 both motivation and confidence are combined as two elements are considered to be inextricably linked (Longmuir et al., 2018). However, there is a need to understand and explore the links and relationship between PL elements, further empirical research such as intervention studies is required to support this. With regards to future tool development, PL elements currently not represented within assessment literature should be a focus to provide assessors with the tools to capture a holistic overview of children's PL and understand more comprehensively the links between elements within the PL concept.

- **Recommendation 4: Longitudinal approaches should be utilised when using summative assessments to capture children's PL**

PL assessment should utilise longitudinal approaches in order to celebrate progress and understand how PL changes over time. Responses from educational stakeholders (study 3, chapter 5) suggested that PL summative assessments should occur frequently within the school year, potentially every school term if possible. Engaging in summative assessments and allowing this to become a part of 'normal school activity' may provide both teachers and children the opportunity to celebrate progress or support for students who are not maintaining a positive relationship with PA (Garder, Harlen and Hayward, 2010). By embracing a longitudinal approach to PL data collection, changes within the PL domain can be assessed in detail, thus providing opportunities to consider areas where students may require additional support while also improving our understanding of how PL may change over time. In

addition, this approach also allows individualised support measures to be utilised as each child's unique PL journey can be considered.

6.4 Formative assessment

Formative assessment has been defined as "activities undertaken by teachers— and by their students in assessing themselves—that provide information to be used as feedback to modify teaching and learning activities" (Black & Wiliam, 2010, p. 82). Therefore, formative assessment may encourage the assessor to utilise a variety of different tools and this process can often be student led (Dixson and Worrell, 2016). Within the UK, the National Foundation for Educational Research report details self-assessment and peer assessment as examples of formative assessment, classifying these as 'informal' assessment styles (NFER, 2007). The purpose of formative assessment is to create a culture where children can make mistakes and learn through receiving descriptive feedback via written comments or verbal conversations. By providing descriptive feedback and co-creating learning goals, progress can be shared with parents or carers outside of the classroom (Black & Wiliam, 2010). In addition, parents/carers are proven to play a crucial role in developing and maintaining healthy lifestyle behaviours such as regular PA and consuming a well-balanced diet (Trost et al., Loprinzi et al., 2010; Shen et al., 2018). Further, parents, carers and guardians already receive feedback on progress in other key subject areas (i.e. mathematics, English) including tests scores, report cards and verbally at parents evening. As such it is important that PL progress is also communicated to encourage PL assessment within and out of school hours. The following recommendations are provided to support teachers, practitioners and children to embrace novel formative methods to assess PL.

- **Recommendation 5: Self-assessment techniques should be utilised to assess PL within primary school children**

Whitehead's recent work suggests that individuals should collect 'data' on their own PL journey in order to inform decisions, set achievable goals and celebrate success (Whitehead, 2019). Further, self-monitoring is a well-known behaviour change technique (Gardner et al., 2015), particularly within the health sector i.e. weight management. In support, a key finding of study 3 (chapter 5) presented self-assessment as a desired and acceptable means of assessment. Thus, self-assessment has been established across stakeholder groups as a feasible and appropriate means of assessing PL and research suggests that children's involvement within the assessment process is crucial for creating opportunities for development and enhancing learning (Lopez-Pastor, Kirk, McPhail and Macdonald, 2013). Specifically, a review of international literature found that through adopting a focus on formative and shared assessment there was clear evidence of improvement across: learning, self-regulation, reliability of students' self-assessment and student satisfaction (Lo´pez-Pastor, 1999, 2006; 2011). Furthermore, Lorente (2005, 2008) and Lorente and Joven (2009) reported consistent results in a longitudinal case study investigating practices that encouraged students to take responsibility for their own learning. Within PL assessment, teachers could prompt an open discussion with students, mutually agreeing on the assessment methods and key elements to focus on, e.g. confidence. Teachers can also encourage students to embrace self-assessment by recording their own assessment scores or describing different PL experiences.

Within the PA research discipline self-report questionnaires (or in the case of young children, proxy-report by a parent/carer) have been used previously as a means of summative assessment to measure SB and PA levels in child populations (Atkin et al., 2012; Lubans et al., 2011). However, self- and proxy-report tools are known to be susceptible to recall errors, misrepresentations and social desirability bias (Atkin et al., 2012; Hardy et al., 2013; Loprinzi & Cardinal, 2011). Therefore, there is a need to look at novel self-report measures to

assess PL. For example, the Digitising Children's Data Collection (DCDC) application (app) was recently used within a mixed-method study to allow flexible data collection with primary school aged children via tablets across multiple settings (Hurter et al., in review). The app can be used within diverse settings and to collect data over a longer period of time than is currently possible with traditional self-report questionnaires which would require repeat administration by a researcher. The DCDC app encourages children to self-report their behaviour daily through photos, voice recordings and drawings at home or within a school setting, thus differing from traditional paper based methods that ask children to recall behaviour from the previous week. The powerful aspect of utilising the DCDC app or similar app in terms of PL assessment is that it allows children to take ownership of their own PL journey, recording and sharing progress as a means of formative assessment. Through further involving children in the assessment process, co-creation of PL strategies can be established between teacher and learner (Morley and Van Rossum, 2019).

- **Recommendation 6: Technology should be used where possible to support PL assessment**

As technology continues to advance, studies have demonstrated the effectiveness of utilising apps and mobile devices within assessment in PE (O'Loughlin, Chróinín and O'Grady, 2013; Penney et al., 2012; Morley and Van Rossum, 2019). Additionally, the findings from the focus groups highlight that all groups recommended aspects of technology to support the assessment process. It is reported that use of devices such as touch screen tablets and apps are becoming increasingly prevalent amongst child population (Neumann, 2017). As such, technology is considered to be an essential aspect of PL self-assessment, supported by the notion that technological devices are becoming increasingly more available (Morley and Van Rossum, 2019). In order to facilitate assessment, tablet devices are recommended, as it has

been reported that within the 7-11 age range self-assessment within PE using video provided an engaging experience for students, regardless of ability (O'loughlin, Chróinín and O'Grady, 2013). Additionally, due to the authentic nature of the tasks, using video to record and replay clips during PE provided students with a realistic and inspiring reflection of their performance (Penney et al., 2012). For example, in relation to skill performance, using video to support the assessment process is reported to improve both the skill performance and motivation to participate in PL (O'loughlin, Chróinín and O'Grady, 2013). While Knudson and Morrison (2002) suggests that being able to pause, stop, rewind and play, video in slow motion may be of particular importance to children with lesser understanding and knowledge of movement principles. In addition, teachers have subsequently reported that the ability to pinpoint and isolate clips can aid the quality of demonstrations provided and feedback given (Weir and Connor, 2009). Using video within PE to observe and compare performance in children is also considered to improve teacher subject knowledge, especially for generalised classroom teachers (Graham, Holt/Hale, and Parker, 2013). While technology can be used to support PL assessment it is also important to consider the negative associations with excessive screen time, particularly within the child population (Kremer et al., 2018). As such, any assessment or prescribed screen time should be informed by available health guidelines (Hinkley et al., 2012)

- **Recommendation 7: The results of PL assessments should be shared and utilised to inform best practice within schools**

Assessing children via both summative and formative assessment will allow for a plethora of information to be collected for each child (i.e. motor competence, perceived confidence, motivation towards being active, barriers to PA, favourite activities, knowledge of PA, after school PA activities, PA achievements). As such, a potential solution for storing and sharing

PL data could be utilising the cloud or a secure shared space within the school intranet, this allowing teacher, students and parents to access and contribute to the evidence base both during school hours and at home. While this has previously been suggested and trialled as a means of collating and storing health data (Thilakanathan et al.,2014; Yue et al., 2012), this approach has not yet been utilised within a primary school setting. It is recommended that summative assessment (i.e. CAPL-2 results) follow a similar approach as displayed with health management cloud systems, particularly when considering the potential factors which are detailed extensively; infrastructure, storage, information and development tendency (Hu et al., 2010). For example, if video clips of children completing the CAMSA were uploaded to a secure share-drive, with consent, researchers specialising in fundamental movement skills could assess and score the assessment electronically at a school, local and national level. Thus, providing access to large sample sizes without having the additional burden of administering and recording the assessments in real time. To continue the example, population data of this kind could be disseminated and information regarding fundamental movement skills used to inform policy and practice, thus striving to improve fundamental movement skills on a population level. In addition, data could be potentially used for specific elements such as fundamental movement skills or larger domain such as the affective domain depending on the availability of the data. While in theory this may be an effective use of PL data, as with all types of personal data the ethical guidelines in place must be stringently followed and ultimately the child and parent must give consent for the information to be shared and utilised in this way (Yue at al., 2016).

- **Recommendation 8: Effective support strategies must be put in place for teachers who are administering PL assessment**

Physical literacy continues to be largely misinterpreted by practitioners (Edwards et al., 2019; Durden-Myers and Keegan, 2019). Therefore, improving practitioner knowledge, understanding and pedagogical skills of PL is of particular importance. Within the focus groups conducted in study 3 (Chapter 5), teachers frequently commented on their perceived lack of PL knowledge and confidence in assessing the concept alongside PE in general. Similarly, expert groups also commented on this, suggesting that teachers would need additional support in order to lead assessments. Moreover, this is supported by statistics that confirm that typically, generalist teachers on average only receive approximately 6 hours of PE training during their teacher-training course (Youth Sport Trust, 2016). As such, assessment within the subject area is often dependent on the confidence and ability of the individual class teacher (Morley and Van Rossum, 2019; Harris, Cale and Musson, 2012; Ni Chróinín and Cosgrave, 2013). An example of additional support measures that could be implemented for teachers is developing short, educational PL sessions that can be developed during continuous professional development time. It is recommended that support from existing organisations are used to facilitate this, i.e. teachers utilising resources and online training courses from the International Physical Literacy Association (IPLA) working towards PL accreditation starting with the PL Foundation Course. Moreover, Edwards et al. (2019) reported that running physical literacy workshops within continuous professional development times resulted in an increase in teachers' knowledge of, and operationalization of, physical literacy. Additionally, dedicating time to encourage self-reflection on their own PL journey may also be a meaningful exercise and provide the teachers with relatable examples to share with their class (Durden-Myers and Keegan, 2019). In order to support the professionals at the forefront of delivering physical literacy assessments, teachers need support in understanding the complexity of the concept of physical literacy, the underpinning philosophy and how PL can be assessed within their child groups.

6.5 Prototype of a PL assessment

Included below is an example of a PL assessment tool, created as a visual representation to summarise the key recommendations discussed above. An IPAD© or similar tablet device is used to support the collection of data, which should be a combination of both formative (child-led) and summative (teacher-led) assessments. Within every class each child, teacher and parent/carer has access to each child's PL folder which includes information regarding the results of assessments, future goals and progress to date. Teachers and parent/carer can review summative assessment score and track progress while children and parents can upload selfies or physical activity experiences to support the ongoing formative assessment process. In addition, teachers can upload the scores from summative assessments and track progress over time. If permission is granted (parent and child consent) the results of summative assessments can be shared with research teams striving to understand the PL concept more comprehensively. Ultimately assessing PL should aim to be a fun and enriching experience for children aged 7-11 years, promoting lifelong engagement in PA.

Figure 1.7 Example of a prototype PL assessment

6.6 Conclusion

The aim of this chapter was to provide useful and pragmatic recommendations for a PL

assessment tool for use within children aged 7-11 years. As such, information from previous

chapters provided the evidence base underpinning these recommendations. Specifically, this

relates to study 2 (chapter 3: systematic review) and study 3 (chapter 4: focus groups),

alongside relevant literature within the research discipline. In particular, this chapter

highlights the importance of using existing tools to assess PL, while embracing both

formative and summative assessment techniques ensuring that all children are provided with

a positive and enriching assessment experience. Key recommendations include ensuring that

assessment experience is fun and utilising technology to ease the burden of assessment for

teachers while supporting engagement in self-assessment within this population. In addition,

the potential to share data sets or profiles electronically should also be explored further to

enable the results of PL assessment to be shared, providing evidence to support population

level research to inform both policy and practice. However, while various recommendations

have been outlined within this chapter, it is important to note that these are at this stage, based on theoretical knowledge rather than practical experience. It is therefore, crucial to utilise this guidance in practice to ascertain essential factors for consideration when assessing PL within the 7-11 age range.

Chapter 7
Synthesis

7.1 Thesis map

Study	Objectives
Study 1: Global interpretations of PL	• To collate, compare, and critically review existing definitions of PL from leading organisations implementing PL agendas around the world **Key findings:** • Seven international groups were identified as working to promote and develop PL, each operating with at least one identifiable definition of the concept • There is a lack of consensus and clarity between definitions and interpretation, as such transparency is needed when reporting alignment to the varying PL concepts • Whitehead's definition (2001) of PL has been utilised or expanded upon within the differing interpretations, providing the rationale for anchoring the work included within this thesis to Whiteheads 'conceptualisation of PL
Study 2: A systematic review of assessments related to the physical domain of PL among children aged 7-11 years	• To systematically review the academic literature for tools to assess the physical domain of PL within children aged 7-11 years. • To explore and critically discuss each assessment tool within the physical, affective and cognitive domains to appraise its (a) psychometric properties; (b) feasibility for use within a primary school setting and (c) alignment to the PL concept. **Key findings:** • There are currently assessments available to assess elements within the physical, cognitive and affective domains of PL • Generally, the psychometric properties of tools were not well reported within the literature, more attention, in particular needs to be given to PROM development and content validity within the field of PA research. • PL assessments are typically proven to be feasible within a school context, however, further research is needed to consider the feasibility of the scoring and administration of assessment tools by teachers as oppose to research teams.
Study 3: Stakeholder perceptions of a PL assessment for children aged 7-11 years	• To explore key stakeholders (academics/practitioners, teachers, and children) views of current practice, future directions and effective implementation of PL assessment, through concurrent focus groups • To inform the development of a rigorous, authentic, and feasible PL assessment for use with children aged 7-11

	Key findings: • Stakeholder groups recognise the demand for a PL assessment tool in order to support the PL concept and inform best practice. • Children valued the perceived enjoyment of assessments and recommended that these mirror positive experiences already occurring within the physical education sector. • Across all stakeholder groups self-assessment was suggested as a suitable method of assessing PL, supported by technology.
Chapter 6: Recommendations for the development of a PL assessment tool for use within children aged 7-11 years	• To outline clear recommendations for a PL assessment within the 7-11 age range • To collate and discuss the findings from previous chapters in detail to provide evidence-based recommendations • To outline assessment resources currently available, while highlighting potential areas for future tool development. **Key findings:** • Where possible utilising technology should be used to support PL assessments, this providing an engaging assessment experience while easing the burden of PL assessment for teachers and practitioners • Creating a positive assessment experience is fundamental to supporting continued engagement in PA and as such should be a crucial factor in the implementation of any PL assessment • The potential to share data sets or profiles electronically should be considered to enable the results of PL assessment shared, providing evidence to support population level research informing both policy and practice.

7.2 Aims and Objectives

The aim of this thesis was to investigate PL assessment within the 7-11 age range. The objectives of the thesis were to a) collate and discuss the varying definitions and interpretations of the concept globally; b) identify and appraise existing assessments for their alignment to the PL concept, methodological rigour and feasibility for use within a primary school setting; c) explore key stakeholders perceptions of PL assessment and assessment in PE ; and d) Provide pragmatic recommendations to inform the assessment of PL within children aged 7-11 years within a primary school setting.

7.3 Summary of Findings

The key findings of each study have been reviewed briefly within the thesis map presented above, this building throughout the development of this thesis. The aim of this section is to review each key finding in depth and outline the implications, which will inform the proposals for the direction of future research.

Study 1:

PL is a concept that has been embraced internationally and as such various misconceptions regarding the definitions of PL have arisen in research, policy and practice (Edwards et al., 2016). The work included within study 1 explored in detail, the multiple definitions of PL currently being utilised globally in order to provide a contemporary update of the concept (Shearer et al., 2018). Collating and comparing PL interpretations provides a much-needed contribution to the current PL research field. Prior to this publication, different definitions had been discussed within the literature (Mandigo et al., 2009; Edwards et al., 2016) but not compared and collated with the purpose of providing clarity. The first key finding within study 1 highlighted that seven groups were identified internationally as currently working to

promote and develop PL, each operating with at least one identifiable definition. The groups included research teams, government organisations (national or state), not-for-profit and corporate groups, or multi-sector partnerships spanning all of these. Typically, each organisations/group used online platforms to share research, present definitions and interpretations. While it is positive to note that PL has been embraced globally, it is also imperative to consider the limitations of different countries working from various interpretations of the concept (Edwards et al., 2017; Shearer et al., 2018). Although the cultural contexts of each individual nation should be considered, communicating the results of assessments and sharing these to inform best practice will be challenging if researchers are assessing fundamentally different concepts (Dudley, 2018). Therefore, an important implication from the results included within study 1, is that a consensus is required to guide future global assessment strategies. Similar to the consensus document produced by Canadian Sport for Life regarding the definition of PL (Tremblay et al., 2018), this level of agreeance and communication is necessary in order to progress the concept further and guide future assessment strategies. While it may be important for countries to adopt their cultural individual identity in regard to the PL concept, consistent, achievable and measurable standards should be agreed upon to successfully guide collaborative research of the PL concept.

Study 2:

There is a considerable lack of evidence relating directly to the assessment of PL in children (Cairney at al., 2019; Edwards et al., 2018). As such, the systematic review (presented in chapter 4) will contribute greatly to the limited evidence base. The review findings report that there are currently 60 different assessments available within the 7-11 age range to measure

the sub-elements within the physical (n=21), cognitive (n=6) and affective domains (n=33) of PL. As such, the availability of assessment tools is an impactful finding from this work as it has highlighted that while there is not one tool to assess the concept in its entirety, there are many different assessments available. Within each PL domain it was found that numerous assessments reported strong methodological rigour (i.e., CAPL-2, TGMD-3, GSPA and PMCS), aligned to Whiteheads interpretation of the concept (i.e., CAMSA, ASK-KIDS, MUGI and ATCPE) and would be feasible for use within a primary school context (i.e., CAPL-2, CAMSA, FG and RCS).

While previous research has outlined potential methods for assessing PL (Edwards et al., 2018), this study was the first to consider not only the availability of assessments but the quality of existing tools. Furthermore, the study provides a plethora of information regarding each existing assessment that can accurately inform and guide the tool selection process- such a detailed insight into each individual tool has not previously been presented within the literature. As such, it is thought that the research included within the systematic review will have a significant impact on the PL research field and encourage assessment through utilising existing tools. Recent research suggests that due to the large and complex nature of PL, assessing aspects of the concept may be a useful approach as opposed to attempting to capture the concept in its entirety (Barnett et al., 2019). Therefore, the work presented within the systematic review will inform this process greatly as teachers, researchers or practitioners can select the most appropriate instrument/s to suit their specific needs.

Another key finding reported within the systematic review is that, in general, the psychometric properties of tools were not well reported within the literature. In fact, studies typically reported only reliability or validity information, with no studies presenting information to satisfy all of the previously presented COSMIN guidelines. The lack of detailed information available regarding the psychometric properties of each assessment

made the quality appraisal stages of the systematic review challenging. Moreover, tool selection becomes increasingly difficult for tool users if this information is not available and this prevents tool users from having confidence in the results of available assessments (Welk, Corbin and Dale, 2000; Mokkink et al., 2018; Griffiths, Toovey, Morgan and Spittle, 2018) . As such, an implication from this project is that any information regarding the psychometric properties of assessments is reported in a transparent manner and the dissemination of such information is considered to be a priority within research teams. Ultimately, without access to this information, decisions regarding the quality of assessments cannot be accurately made, this preventing the successful assessment of PL within children.

Feasibility is an essential factor to consider within the tool development process (Robertson et al., 2017), particularly to consider if an assessment can be utilised within a primary school setting (Klingberg et al., 2019). The results of the systematic review were generally positive in regards to the feasibility of conducting assessments across each domain within a school setting. Specifically, for assessments within the physical domain the time to complete, equipment required and space needed for the majority of assessments suggest that a primary school is an appropriate setting to complete physical competence assessments. However, within both the cognitive and affective domains crucial information at times was not typically reported, i.e. time, space and equipment needed, training required to conduct assessments. While it may be a reasonable assumption that these would be positive due to the majority of assessments being paper based questionnaires it is important that this information is reported in a clear and transparent manner. Similar to the issues discussed above with the inconsistent reporting of psychometric properties, without access to this information decisions regarding the quality of assessments cannot be accurately made, this preventing the successful assessment of PL within children. A key implication from the findings reported within this chapter suggest that reporting all relevant information in regards to

methodological rigour, feasibility and PL alignment in a clear and transparent manner should be a priority for research teams going forward.

Study 3:

While it is often considered that expert opinion is crucial to informing the developmental process of an assessment tool, particularly at the design phase (Longmuir, et al., 2018; Barnett et al., 2019; Mokkink et al., 2018; Robertson et al., 2018) within a primary school context arguably the teachers, support staff and children possess the expertise. Similar to previous studies in related areas (Morley, Van Rossum, Richardson & Foweather, 2018; Ni Chróinín & Cosgrave, 2013), study 3 used qualitative techniques to empower relevant stakeholders to play an active role in the tool development and selection process (Baum, MacDougall and Smith, 2006). All stakeholder groups recognised the demand for a PL assessment tool in order to support the PL concept and inform best practice. This was a positive finding as generally all stakeholders acknowledged the importance and value of the PL concept. Furthermore, stakeholders agreed that having an individualised and person-centred approach may be the best means of assessing the PL concept. The work included within in this study is of particular importance as it is one of few studies within the PL field that has given stakeholders an opportunity to share their experience and perceptions of the concept. Additionally, person-centred approaches have been recommended by Whitehead (2019) as a method of aligning with the PL philosophy, therefore including children within the discussion surrounding assessment was a novel and powerful approach. The results of the focus groups found that children valued the perceived enjoyment of assessments and recommended that these mirror positive experiences already occurring within the physical education sector. As such, providing positive and enriching experiences in PA contexts should be the goal of all PL practices and assessors should ensure that this continues into the assessment experience. Whitehead (2019) recently provided practical guidance for aligning

with the philosophy of the PL concept and this predominantly focused on the assessment. While there is currently research detailing the consequences of a negative experience in PE (Maloney, Beilock and Levine, 2018), there is currently limited practical guidance available for creating positive experiences specifically within assessment scenarios. Furthermore, across all stakeholder groups' self-assessment was suggested as a suitable method of assessing PL, supported by technology. While this in theory is an appropriate method of assessing PL, further research into the practicalities of using technology is required particularly, within the 7-11 age range. Furthermore, while the feasibility of conducting assessments within a primary school setting was considered likely to be positive, it is also important to consider the challenges that may be presented if children are using self-assessment methods supported by teachers (Atkin et al., 2012; Hardy et al., 2013; Loprinzi & Cardinal, 2011).

The results of the research outlined in this thesis highlight that there is a significant lack of empirical research available within the PL discipline particularly in terms of assessment. While the PL concept continues to grow and be implemented within national policies, it is expected that the body of research available will also increase. It is suggested that researchers, practitioners, teachers and children start assessing PL using available tools and utilise the results to provide further clarity and opportunities for learning within the discipline. PL is a concept that captures engagement in PA across the life course, as such assessment while starting in the child population should strive to continue to support people on their PL journey.

Recommendations chapter:

The purpose of this chapter was to provide pragmatic recommendations to enable PL assessment. In particular, this chapter highlights the importance of using existing tools to

assess PL, while embracing both formative and summative assessment techniques, ensuring that all children are provided with a positive and enriching assessment experience. The recommendations within this chapter are predominantly derived from the findings of each study within the thesis and have provided a guide for teachers and practitioners who are considering assessing PL within the 7-11 age range. This chapter provides clear worked examples alongside a 'PL prototype' while using terminology/constructs already well known to the individuals working within the primary school setting, this creating a novel and useful approach. Key recommendations discussed within this chapter include; ensuring that assessment experience is fun, utilising technology where possible to ease the burden of assessment and exploring the utilisation of self-assessment techniques. In addition, a combination of both formative and summative assessment techniques should be used with the potential to share data sets or profiles electronically should be considered to enable the results of PL assessment to be shared, this providing evidence to support population level research to inform both policy and practice.

7.4 Strengths and limitations

Study 1 (chapter 3) collated and compared the varying definitions and international interpretations of PL, disseminating this in an accessible manner is particularly valuable as it can be used to guide future research. Specifically, as it is recommended that both researchers and practitioners should clearly outline any definitions of PL used within research and practice (Edwards et al., 2016). Study 1 embraced a narrative review technique which enabled a comprehensive and critical analysis of PL definitions used internationally. A key strength of the review was the wide scope, specifically, the literature published in books, electronic/paper-based journal articles and publically available websites was included. This ensured that a thorough representation of each definition was explored, alongside information

that was being shared openly within the public. However, the organisations included were limited to prominent, English speaking research groups working with at least one identifiable definition of PL. In addition, the study provided a comparison of definitions but lacked overall detail on how each definition was operationalised within each individual context. Further, while sections of the study were contributed by authors within the relevant context (i.e. Lowri Edwards, Wales) the majority of work was conducted by two researchers working and living in the UK (CS and HG). As such, a strength of future research would be to commission representatives from each country to give an overview of their national perspective, this providing the necessary detail within policy and practice. Study 1 provides a clear starting point for understanding the direction and comprehension of PL internationally, however it is necessary to understand global interpretations of PL in greater detail from the perspective of individuals working within the relevant organisations.

The systematic review, presented in study 2 is the first of its kind to appraise existing tools available within each PL domain with regards to alignment to the PL concept, methodological rigour and the feasibility of each tool. As such, this has provided an extensive and detailed information base regarding each assessment tool which has the potential to greatly inform PL assessment within the 7-11 age range. In addition, the review adhered to a systematic and transparent methodology which has enabled robust research findings to be produced. Specifically, the methodological quality of each included assessment tool was appraised using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) risk of bias checklist (Terwee et al., 209; Mokkink et al., 2018). The COSMIN checklist is highly regarded as it was developed in a multidisciplinary, international consensus-study in which 43 experts in health status measurement participated internationally (Mokkink et al., 2012) as such, using COSMIN is considered to be a key strength of this work. Further, transparency is a crucial element of conducting systematic reviews (Khan,

Kunz, Kleijnen and Antes, 2003) and due to the detailed instruction manual and the background information provided on the development and validation of the checklist this ensured the relevant stages within the review adhered to stringent guidelines, with results recorded in a clear, transparent manner (Terwee et al., 209; Mokkink et al., 2018). The feasibility, interpretability, cost and ease of use of each instrument was reviewed using a utility matrix based upon work from Klingberg et al (2018). The checklist was originally developed with primary schools as a key contextual setting and considered generalist teachers as the assessor which is consistent with setting and age range within this study. Again, the addition of another lens in which to appraise each PL assessment tool provides a practical and relevant means of guiding PL assessment. In particular, the addition of feasibility appraisal can support teachers in selecting an appropriate tool for their individual setting. This is an importation addition of this work and a crucial aspect of the development phase within tool development that is often overlooked (Barentt et al., 2019). To complete the appraisal of each tool, a bespoke PL checklist was developed to ascertain alignment to the concept, a novel aspect of this work. The checklist was developed from work presented by Whitehead (2010) and the Australian Spots Commission (ACS) (2017) and this is the first study to map the PL sub-elements in this manner. It is important to also note that within study 2, the various rounds of screening, data collection and quality appraisal of tools were conducted independently by two individual reviewers (CS and HG). This is in line with PRISMA-P guidelines (Moher et al., 2015) and improved the reliability of results and overall rigour of the review process. Furthermore, the protocol information for this review, was registered with PROSPERO (reference **CRD42017062217**) as means of further ensuring transparency of the methods.

Conducting research within a discipline that is in its infancy stages is a complex process and therefore the work included within study 2 had various limitations alongside the

many strengths. Firstly, the lack of consistency and clarity across global interpretations of the PL concept proved to be problematic (see study 1) as assessing a concept that is not definitely defined is a challenging feat. Once the full extent of the differences internationally were realised, various working group meetings had to be scheduled with the research team in order to decide on search terms and domains of inclusion for the systematic review (study 2: chapter 4), i.e. exclusion of the social domain. The exclusion of a represented domain within the PL concept may have affected overall findings of the systematic review as the search terms included only Whitehead's and the Australian Sports Commission interpretation of the concept. Furthermore, only studies published in peer review journals and in the English language were accepted into the review meaning that some existing assessments may not have been represented within this work i.e. assessment used in practice, assessments not published in English. Furthermore, only assessments within the 7-11 age range were included within the review meaning that assessment in older or younger populations maybe available to be adapted for use with the 7-11 age range however this was outside the scope of the review.

Study 3 reported the findings from 26 different focus groups conducted concurrently with a range children, teachers and self-defined PL experts. It is typical for a study to present the findings of between 4-6 focus groups (Patton, 2005), therefore, to include such a large number of groups is considered to be a key strength of this study. Also, as the groups were conducted concurrently a detailed snapshot of a specific time period is captured. In addition, including children within the focus groups is a novel and important approach as it enabled children to share their experiences and perspectives, this proving relevant information regarding their individual context (Morely, Van Rossum, Richardson and Foweather, 2019). This approach also closely aligns to Whiteheads underpinning philosophies of phenomenology and existentialism (Whitehead, 2019). The data collected related to focus

groups conducted in England, Wales and Scotland, in a range of demographic settings, this providing a powerful insight into different educational and academic contexts. As perceptions of PL are generally unknown within both teacher and child groups, this provides a much needed detailed and in-depth exploration (Smith and Sparkes, 2016; Gibson 2007; Domville, 2018). To facilitate the focus groups, a semi-structured interview guide was developed based on recommendations from Bowen et al. (2009) who suggested several areas of focus when exploring feasibility of new assessments. Recommendations from Bowen et al (2009) are commonly used within feasibility studies, typically when considering the feasibility for practitioners (Lander et al., 2016; Klingberg et al., 2018). In addition, the framework provided clear direction that sufficiently assisted with the exploration of research questions within this thesis, another important strength. Furthermore, each focus group was transcribed and analysed both deductively (using Bowen et al. (2009) as a thematic framework) and inductively, enabling additional themes to be generated (Braun & Clarke, 2006; 2019). Similar to previous studies in related areas (Morley, Van Rossum, Richardson & Foweather, 2018; Ni Chróinín & Cosgrave, 2013), the use of focus groups allowed for deep and meaningful insight into the perspectives of participants, which subsequently allowed for the construction of meaningful themes. The findings however do not necessarily allow for generalisation for different age groups, or assessments conducted outside of the school environment and this is a potential limitation of this work. Moreover, it is important to consider potential sampling bias (Tuckett et al., 2004) within participants i.e. as inclusion in the study was voluntary, schools who responded to the invitation to take part may have been more likely to support and value PL and PE practices, this affecting the results. In addition, in-depth analysis across the United Kingdom was outside of the scope of study 3 and this could be considered as a starting point for future research. Further, study 3 explored stakeholder perceptions of PL assessment (teachers, children and self-defined PL experts),

parents were not included and are seen as a key influencer of promoting positive PA practices of children within this age range (Trost et al., 2003; Edwardson and Gorley, 2010; Wilk et al., 2018). Therefore, it is important that parents are included within the conversation moving forward in order to ascertain how PL assessment can be supported beyond the school day. Finally, while the concept of PL encourages assessment throughout the life course i.e. beyond school PE and across education, sporting, recreation, and health contexts (Barnett et al., 2019), this was also outside the scope of the current project, and further research across the aforementioned areas is encouraged.

7.5 Originality

The project of work included within this thesis provides a unique and in-depth exploration of PL assessment in children. While researchers, practitioners and policy makers have called for a need for a PL assessment tool, the findings from this project highlight the need to firstly consider and potentially utilise existing assessments to understand the PL concept further. The findings suggest an alternative direction for PL assessment, whereby PL indicators or select elements can be assessed in the first instance to guide the implementation of future support strategies (Barnett, et al., 2019). Furthermore, a novel aspect of this research is that the voices of teachers, children and stakeholders are prominent throughout the findings. By engaging with and respecting stakeholder voices an in-depth and powerful representation of PL assessment perspectives was captured through the qualitative research conducted within this thesis.

7.6 Recommendations for future research, policy and practice

Based on the findings presented in this thesis there are several recommendations to further the research and practice of PL assessment in children within the 7-11 age range:

- Research into PL assessment should continue, particularly within the 7-11 age range. Cross-sectional studies, case studies, cohort studies and intervention studies in particular should be used to understand PL within the 7-11 child population and should strive to include at least one PL assessment (i.e., CAPL-2).

- Stakeholders should continue to be included within the PL assessment conversation and future assessment decisions in order to provide assessment strategies that are relevant and useful within the primary school context

- The results of PL assessments should be disseminated widely (via academic and practitioner channels) with the aim of supporting population level research, ongoing policy development and informing 'best' practice in regards to PL assessment

- PL assessment should strive to align closely to Whiteheads (2019) underpinning philosophy; particular with regards to promoting an individualised and positive assessment experience

- Teachers, children and practitioners should be encouraged to experiment with different forms of PL assessment (both summative and formative) to explore the feasibility of using assessments within a typical school day

- Research into future PL tool development should conduct a range of psychometric testing in line with the COSMIN guidelines and strive to report these in a transparent and publicly available manner (i.e. published, included within testing manual, COSMIN website)

- The PL checklist provided within this thesis can be used to guide further research investigating the links between PL elements. Particular focus should be given to elements that are linked to a range of sub-elements and how best to support the assessment of these in practice

- PL encourages engagement in PA across the life course and as such assessment across different ages and stages of life is required. In particular research should focus on existing assessment methods (i.e. self-monitoring) and consider how these may be supported throughout an individual's lifetime

- Assessments that are currently being used in practice (i.e. not published via academic channels) must be explored further to ascertain if they can be used to support PL assessment in children within the 7-11 age range

- Self-assessment practices should be utilised within the 7-11 age as a means of formative PL assessment, the practicalities of this should also be explored further through intervention and case studies with the results being widely disseminated to inform future PL assessment strategies

- PL training courses should be made available to ALL teachers and practitioners either through CPD or initial teacher training courses. Similar to other core subjects within the curriculum, assessment in PE should be a foci of training with the aim of empowering teachers to support effective PL assessment within their pupils

- PL assessment should strive to include parents, carers, friends and family and be conducted both within school and out of school hours. In particular schools should consider best to support this i.e. opening school playgrounds and facilities to promote PA assessment out-of-school hours or during school holiday periods

- PL should continue to be included within health and education policies, however further detail and resources are required to support assessment of PL in practice across a range of sectors i.e. health, leisure, education and recreation

- The Department of Health recently updated the PA guidelines for children, adults and the elderly. It is recommended that PL guidelines should also be made publicly

available with opportunities to complement the existing PA guidelines being explored

fully

7.7 Reflection

From both a personal and professional perspective, the PhD process outlined within this

thesis has been an ongoing learning and developmental experience. My knowledge base and

comprehension of the PL concept has increased exponentially during the course of this

project. In addition, my research expertise has expanded tremendously through use of a range

of quantitative and qualitative research methodologies. These include but are not limited to

reviewing international literature, systematic reviewing and facilitating focus groups.

Throughout this project I have had the pleasure of working with participants from all walks

of life, each at different stages of their PL journey, this in particular has inspired me to

continue to work with this field. The project has given me an invaluable opportunity to reflect

on my own relationship with PA and consider the progress I am making within my own PL

journey. Through connecting with the underpinning philosophy of the PL concept this

experience has left me not only enlightened but passionate about supporting others to develop

and maintain positive relationships with PA.

7.8 Conclusions

This thesis has provided a unique exploration of PL assessment within the 7-11 age range.

Not only did the work included within this thesis present assessments but the research also

provided opportunities for key stakeholders to engage and share their experiences of PL

assessments. This thesis has highlighted that while there are assessment tools currently

available to measure elements of the PL concept more work is required to explore assessment

of the concept holistically and in line with Whiteheads interpretation and philosophical

underpinnings. These findings demonstrate a clear need for further research within the

discipline, focusing on assessment of PL in child populations. This work has attempted to consider the assessment culture within the UK and provide recommendations that would be feasible within this setting and it is hoped that these recommendations are taken forward to support the assessment of PL within the 7-11 age range. Research into the PL concept is in its infancy stages and this work will be a timely addition to support the assessment of PL within children, aiming to engage and inspire children to be active for life. Future research should look to examine the development of PL interventions that can be implemented at the primary school stage of development, in order to support children on their PL journey to progress on to leading physically active lives.

References

Angell, C., Alexander, J., & Hunt, J. A. (2015). 'Draw, write and tell': A literature review and methodological development on the 'draw and write' research method. Journal of Young children Research, 13(1), 17-28.

Allan, V., Turnnidge, J., & Côté, J. (2017). Evaluating approaches to physical literacy through the lens of positive youth development. Quest, 69(4), 515-530.

Almond, L. (2013). What is the value of physical literacy and why is physical literacy valuable?. Presidents Message 12.

Araujo, D. and Davids, K., 2009. Ecological approaches to cognition and action in sport and exercise: Ask not only what you do, but where you do it. International Journal of Sport Psychology, 40(1), p.5.

Araújo-Soares, V., Sniehotta, F. F., Laing, C. M., Gellert, P., Jackson, D., & Speakman, J. R. (2015). Social cognitions measured in 4 to 6 year olds are predictive of objectively measured physical activity. Psychology & health, 30(10), 1240-1257.

Armour, K., Quennerstedt, M., Chambers, F., & Makopoulou, K. (2017). What is 'effective'CPD for contemporary physical education teachers? A Deweyan framework. Sport, education and society, 22(7), 799-811.

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2016) Health and physical education (HPE). Canberra, ACT: Commonwealth of Australia. Retrieved from https://www.australiancurriculum.edu.au/f-10-curriculum/health-and-physical-education/

Australian Sports Commission (2017). Draft Australian physical literacy standard. Retrieved from http://ausport.gov.au/physical_literacy

Babic, M. J., Morgan, P. J., Plotnikoff, R. C., Lonsdale, C., White, R. L., & Lubans, D. R. (2014). Physical activity and physical self-concept in youth: systematic review and meta-analysis. Sports medicine, 44(11), 1589-1601.

Bailey, R., Armour, K., Kirk, D., Jess, M., Pickup, I., Sandford, R., & Education, B. P. (2009). The educational benefits claimed for physical education and school sport: an academic review. Research papers in education, 24(1), 1-27.

Bandelli, L. N., Gray, H. L., Paul, R. C., Contento, I. R., & Koch, P. A. (2017). Associations among measures of energy balance related behaviors and psychosocial determinants in urban upper elementary school children. Appetite, 108, 171-182.

Banville, D., Dyson, B., Kulinna, P. H., & Stylianou, M. (2019). Classroom teachers' and administrators' views of teaching health and physical education. *European Physical Education Review*, 1356336X19867731.

Barnett, L. M., Hardy, L. L., Brian, A. S., & Robertson, S. (2015). The development and validation of a golf swing and putt skill assessment for children. *Journal of sports science & medicine*, *14*(1), 147.

Barnett, L. M., Ridgers, N. D., Zask, A., & Salmon, J. (2015). Face validity and reliability of a pictorial instrument for assessing fundamental movement skill perceived competence in young children. Journal of science and medicine in sport, 18(1), 98-102.

Barnett, L. M., Stodden, D., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., ... & Lander, N. J. (2016). Fundamental movement skills: An important focus. *Journal of Teaching in Physical Education*, *35*(3), 219-225.

Barnett, L. M., Dudley, D. A., Telford, R. D., Lubans, D. R., Bryant, A. S., Roberts, W. M., ... & Salmon, 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.

Barnett, L. M., Vazou, S., Abbott, G., Bowe, S. J., Robinson, L. E., Ridgers, N. D., & Salmon, J. (2016). Construct validity of the pictorial scale of Perceived Movement Skill Competence. Psychology of Sport and Exercise, 22, 294-302.

Beattie, M., Lauder, W., Atherton, I., & Murphy, D. J. (2014). Instruments to measure patient experience of health care quality in hospitals: a systematic review protocol. *Systematic reviews*, *3*(1), 4.

Beattie, M., Murphy, D. J., Atherton, I., & Lauder, W. (2015). Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. *Systematic reviews*, *4*(1), 97.

Belanger, K., Barnes, J. D., Longmuir, P. E., Anderson, K. D., Bruner, B., Copeland, J. L., ... & Law, B. (2018). The relationship between physical literacy scores and adherence to Canadian physical activity and sedentary behaviour guidelines. *BMC public health*, *18*(2), 1042.

Beni, S., Fletcher, T., & Ní Chróinín, D. (2017). Meaningful experiences in physical education and youth sport: A review of the literature. *Quest*, *69*(3), 291-312.

Beni, S., Ní Chróinín, D., & Fletcher, T. (2019). A focus on the how of meaningful physical education in primary schools. Sport, Education and Society, 24(6), 624-637.

Biddle, S. J., & Asare, M. (2011). Physical activity and mental health in children and adolescents: a review of reviews. *British journal of sports medicine*, *45*(11), 886-895.

Bingham, D. D., Costa, S., Hinkley, T., Shire, K. A., Clemes, S. A., & Barber, S. E. (2016). Physical activity during the early years: a systematic review of correlates and determinants. American journal of preventive medicine, 51(3), 384-402.

Black, P. (2015). Formative assessment–an optimistic but incomplete vision. Assessment in Education: Principles, Policy & Practice, 22(1), 161-177.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education), 21(1), 5.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. Phi delta kappan, 86(1), 8-21.

Blair, R., & Capel, S. (2008). The use of coaches to cover planning, preparation and assessment time–some issues. *Primary Physical Education Matters*, *3*(2), 9-10.

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., ... & Fernandez, M. (2009). How we design feasibility studies. *American journal of preventive medicine*, *36*(5), 452-457.

Booth, A., Moore, G., Flemming, K., Garside, R., Rollins, N., Tunçalp, Ö., & Noyes, J. (2019). Taking account of context in systematic reviews and guidelines considering a complexity perspective. *BMJ global health*, *4*(Suppl 1), e000840.

Bornholt, L. J., & Piccolo, A. (2005). Individuality, Belonging, and Children's Self Concepts: A Motivational Spiral Model of Self-Evaluation, Performance, and Participation in Physical Activities. *Applied Psychology*, *54*(4), 515-536.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589-597.

Brown, G. T., Andrade, H. L., & Chen, F. (2015). Accuracy in student self-assessment: directions and cautions for research. Assessment in Education: Principles, Policy & Practice, 22(4), 444-457.

Brown, G., & Harris, L. R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency.

Browne, T. (2015). A case study of student teachers' learning and perceptions when using tablet applications teaching physical education. Asia-Pacific Journal of Health, Sport and Physical Education, 6(1), 3-22.

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.

Cairney, J., Dudley, D., Kwan, M., Bulten, R., & Kriellaars, D. (2019). Physical literacy, physical activity and health: Toward an evidence-informed conceptual model. Sports Medicine, 49(3), 371-383.

Cairney, J., Veldhuizen, S., Graham, J. D., Rodriguez, C., Bedard, C., Bremer, E., & Kriellaars, D. (2018). A Construct Validation Study of PLAYfun. *Medicine and science in sports and exercise*, *50*(4), 855-862.

Cairney, J., Kiez, T., Roetert, E. P., & Kriellaars, D. (2019). A 20th-Century Narrative on the Origins of the Physical Literacy Construct. *Journal of Teaching in Physical Education*, *38*(2), 79-83.

Calatayud, J., Martín, F., Colado, J. C., Benavent, J., Martínez, M. C., & Flández, J. (2017). Relationship between the modified star excursion balance test and the 4x10 m shuttle run test in children. *Cultura, Ciencia y Deporte*, *12*(35), 111-116.

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.

Calfas, K. J., Sallis, J. F., & Nader, P. R. (1991). The development of scales to measure knowledge and preference for diet and physical activity behavior in 4-to 8-year-old children. *Journal of Developmental and Behavioral Pediatrics*.

Canadian Sport For Life (CS4L). (2015). Canada's physical literacy consensus statement. Retrieved from: http://physicalliteracy.ca/physical-literacy/consensus-statement/

Canadian Sport For Life (CS4L). (2017). Physical literacy definition. Retrieved from http://physicalliteracy.ca/physical-literacy/

Carse, N., Jess, M., & Keay, J. (2018). Primary physical education: Shifting perspectives to move forwards. *European Physical Education Review*, *24*(4), 487-502.

Carson, V., Tremblay, M. S., Chaput, J. P., & Chastin, S. F. (2016). Associations between sleep duration, sedentary time, physical activity, and health indicators among Canadian children and youth using compositional analyses. Applied Physiology, Nutrition, and Metabolism, 41(6), S294-S302.

Casey, A., Goodyear, V. A., & Armour, K. M. (2017). Rethinking the relationship between pedagogy, technology and learning in health and physical education. Sport, education and society, 22(2), 288-304.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public health reports, 100(2), 126.

Castelli, D. M., Barcelona, J. M., & Bryant, L. (2015). Contextualizing physical literacy in the school environment: The challenges. *Journal of Sport and Health Science*, *4*(2), 156-163.

Castelli, D. M., Centeio, E. E., Beighle, A. E., Carson, R. L., & Nicksic, H. M. (2014). Physical literacy and comprehensive school physical activity programs. Preventive medicine, 66, 95-100.

Castelli, D. M., & Chen, A. (2018). Large-scale physical education interventions: past, present, and future. Kinesiology Review, 7(3), 259-265.

Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. Educational research review, 14, 33-46.

Chen, A. (2015). Operationalizing physical literacy for learners: Embodying the motivation to move. Journal of Sport and Health Science, 4, 125-131.

Cosgrove, J. M., Chen, Y. T., & Castelli, D. M. (2018). Physical fitness, grit, school attendance, and academic performance among adolescents. *BioMed research international, 2018.*

Corbin, C. B. (2016). Implications of physical literacy for research and practice: A commentary. *Research quarterly for exercise and sport*, *87*(1), 14-27.

Craxford, S., Deacon, C., Mynt, Y., & Ollivere, B. (2019). Assessing outcome measures used after rib fracture: A COSMIN systematic review. Injury.

Creswell, JW., & Creswell, JD., (2018), Research design; Qualitative, quantitative & mixed approaches (5th ed), Thousand Oaks: Sage

Dennett, D.C. (1995.) Darwin's dangerous idea: Evolution and the meanings of life. New York. NY: Simon and Schuster.

Department for Digital, Culture, Media & Sport., (2015). Sporting future: A new strategy for an active nations. Retrieved from https://www.gov.uk/government/publications/sporting-future-a-new-strategy-for-an-active-nation

Department for Education, (September, 2013). National Curriculum in England: PE programmes of study. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-physical-education-programmes-of-study

Department for Education (March, 2017). Statutory framework for the early years foundation stage: Setting the standards for learning, development and care for children from birth to five. Retrieved from https://www.gov.uk/government/publications/early-years-foundation-stage-framework--2

Department of Health., (2011), Start Active, Stay Active: A report on physical activity from the four home countries' Chief Medical Officers, London: Department of Health and Social Care. Retrieved from https://www.gov.uk/government/publications/start-active-stay-active-a-report-on-physical-activity-from-the-four-home-countries-chief-medical-officers

Department of Health and Social Care., (2016). Childhood obesity: A plan for action. Retrieved from https://www.gov.uk/government/publications/childhood-obesity-a-plan-for-action

Department of Health and Social Care., (2019), Physical activity guidelines: UK Chief Medical Officer's report. Chief Medical Officers, London: Department of Health and Social Care. Retrieved from https://www.gov.uk/government/publications/physical-activity-guidelines-uk-chief-medical-officers-report

Derri, V., Tsapakidou, A., Zachopoulou, E., & Kioumourtzoglou, E. (2001). Effect of a music and movement programme on development of locomotor skills by children 4 to 6 years of age. European Journal of Physical Education, 6(1), 16-25.DinanThompson, M., & Penney, D. (2015). Assessment literacy in primary physical education. *European Physical Education Review*, *21*(4), 485-503.

DinanThompson, M., & Penney, D. (2015). Assessment literacy in primary physical education. European Physical Education Review, 21(4), 485-503.

Domville, M. S., Watson, P. M., Richardson, D. J., & Graves, L. E. (2018). Educator perspectives on factors influencing children's school-based physical activity. *Health promotion international*.

Domville, M., Watson, P. M., Richardson, D., & Graves, L. E. F. (2019). Children's perceptions of factors that influence PE enjoyment: a qualitative investigation. Physical Education and Sport Pedagogy, 24(3), 207-219.

Donaldson, G. (2015). Successful futures: Independent review of curriculum and assessment arrangements in Wales. Retrieved from http://dera.ioe.ac.uk/22165/2/150225-successful-futures-en_Redacted.pdf

Dudley, D. A. (2015). A conceptual model of observed physical literacy. *The Physical Educator*, *72*(5).

Dudley, D., Cairney, J., Wainwright, N., Kriellaars, D., & Mitchell, D. (2017). Critical considerations for physical literacy policy in public health, recreation, sport,

Dyson, B., Cowan, J., Gordon, B., Powell, D., & Shulruf, B. (2018). Physical education in Aotearoa New Zealand primary schools: Teachers' perceptions and policy implications. European Physical Education Review, 24(4), 467-486.

Economos, C. D., Hennessy, E., Sacheck, J. M., Shea, M. K., & Naumova, E. N. (2010). Development and testing of the BONES physical activity survey for young children. *BMC musculoskeletal disorders*, *11*(1), 195.

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.

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.

Edwards, L. C., Bryant, A. S., Morgan, K., Cooper, S. M., Jones, A. M., & Keegan, R. J. (2019). A Professional Development Program to Enhance Primary School Teachers' Knowledge and Operationalization of Physical Literacy. *Journal of Teaching in Physical Education*, *38*(2), 126-135.

Ennis, C. D. (2015). Knowledge, transfer, and innovation in physical literacy curricula. Journal of Sport and Health Science, 4(2), 119-124.

Erbaugh, S. J. (1978). Assessment of swimming performance of preschool children. Perceptual and motor skills, 46(3_suppl2), 1179-1182.

Ericsson, I. (2007). MUGI observation checklist: An alternative to measuring motor skills in physical education classes. Asian Journal of Exercise & Sports Science, 4(1).

Ericsson, I. (2008). To measure and improve motor skills in practice. International Journal of Pediatric Obesity, 3(sup1), 21-27.

España-Romero, V., Artero, E. G., Jimenez-Pavón, D., Cuenca-Garcia, M., Ortega, F. B., Castro-Pinero, J., ... & Ruiz, J. R. (2010). Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA Study. International journal of sports medicine, 31(07), 490-497.

Faigenbaum, A. D., Bagley, J., Boise, S., Farrell, A., Bates, N., & Myer, G. D. (2015). Dynamic balance in children: Performance comparison between two testing devices. Athletic Training and Sports Health Care, 7(4), 160-164.

Fjørtoft, I. (2000). Motor fitness in pre-primary school children: the EUROFIT motor fitness test explored on 5–7-year-old children. Pediatric exercise science, 12(4), 424-436.

Fletcher, T., & Mandigo, J. (2012). The primary schoolteacher and physical education: a review of research and implications for Irish physical education. *Irish Educational Studies*, *31*(3), 363-376.

J. Foulkes, personal communication, May 9 2019.

Foweather, L., Knowles, Z., Ridgers, N. D., O'Dwyer, M. V., Foulkes, J. D., & Stratton, G. (2015). Fundamental movement skills in relation to weekday and weekend physical activity in preschool children. Journal of science and medicine in sport, 18(6), 691-696.

Furtado Jr, O., & Gallagher, J. D. (2012). The reliability of classification decisions for the Furtado-Gallagher computerized observational movement pattern assessment system—FG-COMPASS. *Research quarterly for exercise and sport*, *83*(3), 383-390.

Giblin, S., Collins, D., & Button, C. (2014). Physical literacy: Importance, assessment, and future directions. Sports Medicine, 44, 1177-1184.

Gondivkar, S. M., Gadbail, A. R., Sarode, S. C., Gondivkar, R. S., Yuwanati, M., Sarode, G. S., & Patil, S. (2019). Measurement properties of oral health related patient reported outcome measures in patients with oral cancer: A systematic review using COSMIN checklist. PloS one, 14(6), e0218833.

Goodway, J. D., Ozmun, J. C., & Gallahue, D. L. (2019). Understanding motor development: Infants, children, adolescents, adults. Jones & Bartlett Learning.

Goodyear, V. A., Kerner, C., & Quennerstedt, M. (2019). Young people's uses of wearable healthy lifestyle technologies; surveillance, self-surveillance and resistance. Sport, education and society, 24(3), 212-225.

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.

Griggs, G. (Ed.). (2012). An introduction to primary physical education. Routledge.

Grix, J. (2002). Introducing students to the generic terminology of social research. *Politics*, *22*(3), 175-186.

Gu, X. (2016). Fundamental motor skill, physical activity, and sedentary behavior in socioeconomically disadvantaged kindergarteners. Psychology, health & medicine, 21(7), 871-881.

Haerens, L., Aelterman, N., Vansteenkiste, M., Soenens, B., & Van Petegem, S. (2015). Do perceived autonomy-supportive and controlling teaching relate to physical education

students' motivational experiences through unique pathways? Distinguishing between the bright and dark side of motivation. *Psychology of sport and exercise*, *16*, 26-36.

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., Ekelund, U., & Lancet Physical Activity Series Working Group. (2012). Global physical activity levels: surveillance progress, pitfalls, and prospects. *The lancet*, *380*(9838), 247-257.

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

Harter, S. (1978). Pleasure derived from challenge and the effects of receiving grades on children's difficulty level choices. Child Development, 788-799.

Hassmén, P., Keegan, R., & Piggott, D. (2016). *Rethinking sport and exercise psychology research: Past, present and future*. Springer.

Hay, P.J. and Penney, D. (2013). Assessment in Physical Education: A Sociocultural Perspective. London: Routledge.

Health and Social Care Information Centre. (2016). *Health Survey for England 2016.* Leeds: Health and Social Care information Centre. Retrieved from https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2016

Herrmann, C., Gerlach, E., & Seelig, H. (2015). Development and validation of a test instrument for the assessment of basic motor competencies in primary school. *Measurement in Physical Education and Exercise Science*, *19*(2), 80-90.

Herrmann, C., & Seelig, H. (2017). Structure and profiles of basic motor competencies in the third grade—validation of the test instrument MOBAK-3. Perceptual and motor skills, 124(1), 5-20.

Hilland, T., Ridgers, N., Stratton, G., Knowles, Z., & Fairclough, S. (2018). Origins of perceived physical education ability and worth among English adolescents. European Physical Education Review, 24(2), 165-180.

Hoeboer, J., De Vries, S., Krijger-Hombergen, M., Wormhoudt, R., Drent, A., Krabben, K., & Savelsbergh, G. (2016). Validity of an Athletic Skills Track among 6-to 12-year-old children. Journal of sports sciences, 34(21), 2095-2105.

Hooker, T. (2017). Transforming teachers' formative assessment practices through ePortfolios. Teaching and Teacher Education, 67, 440-453.

Howell, A. (2015). Exploring children's lived experiences of NAPLAN. In *National Testing in Schools* (pp. 164-180). Routledge.

Hulteen, R. M., Lander, N. J., Morgan, P. J., Barnett, L. M., Robertson, S. J., & Lubans, D. R. (2015). Validity and reliability of field-based measures for assessing movement skill competency in lifelong physical activities: a systematic review. Sports Medicine, 45(10), 1443-1454.

Hyndman, B., & Pill, S. (2018). What's in a concept? A Leximancer text mining analysis of physical literacy across the international literature. *European Physical Education Review*, *24*(3), 292-313.

International Physical Literacy Association. (2017). IPLA definition. Retrieved from https://www.physical-literacy.org.uk/

Jacquez, F., Vaughn, L. M., & Wagner, E. (2013). Youth as partners, participants or passive recipients: A review of children and adolescents in community-based participatory research (CBPR). *American journal of community psychology*, *51*(1-2), 176-189.

Jefferies, S. (2016). What exactly is 50 million strong by 2029? - SHAPE America. Retrieved from http://community.shapeamerica.org/blogs/steve-jefferies/2016/04/01/what-exactly-is-50-million-strong-by-2029.

Jess, M., Keay, J., & Carse, N. (2016). Primary physical education: A complex learning journey for children and teachers. *Sport, Education and Society*, *21*(7), 1018-1035.

Jess, M., McEvilly, N., & Carse, N. (2017). Moving primary physical education forward: start at the beginning. *Education 3-13*, *45*(5), 645-657.

Jones, B. A. (1988). A scale to measure the attitudes of school pupils towards their lessons in physical education. Educational Studies, 14(1), 51-63.

Jurbala, P. (2015). What is physical literacy, really?. *Quest*, *67*(4), 367-383.

Keegan, R. J., Barnett, L. M., Dudley, D. A., Telford, R. D., Lubans, D. R., Bryant, A. S., ... & Vella, S. A. (2019). Defining physical literacy for application in Australia: a modified Delphi method. *Journal of Teaching in Physical Education*, *38*(2), 105-118.

Keegan, R. J., Keegan, S. L., Daley, S., Ordway, C., & Edwards, A. (2013). Getting Australia moving: Establishing a physically literate & active nation (GAME PLAN). Retrieved from http://www.canberra.edu.au/researchrepository/file/50f8c79c-2aca-a83f-aee8-254288c36220/1/full_text_final.pdf

Kirk, D. (2009). *Physical education futures*. Routledge.

Kirk, D. (2013). Educational value and models-based practice in physical education. Educational Philosophy and Theory, 45(9), 973-986.

Kitzinger, J. (1995). Qualitative research: introducing focus groups. Bmj, 311(7000), 299-302.

Klingberg, B., Schranz, N., Barnett, L. M., Booth, V., & Ferrar, K. (2018). The feasibility of fundamental movement skill assessments for pre-school aged children. *J. Sports Sci*, *7*, 1-9.

Koekoek, J., & Van Hilvoorde, I. (Eds.). (2018). *Digital technology in physical education: Global perspectives*. Routledge.

Lander, N., Eather, N., Morgan, P. J., Salmon, J., & Barnett, L. M. (2017). Characteristics of teacher training in school-based physical education interventions to improve fundamental movement skills and/or physical activity: A systematic review. *Sports Medicine*, *47*(1), 135-161.

Lakes, K. D. (2013). Measuring self-regulation in a physically active context: Psychometric analyses of scores derived from an observer-rated measure of self-regulation. Mental health and physical activity, 6(3), 189-196.

Lakes, K. D., & Hoyt, W. T. (2004). Promoting self-regulation through school-based martial arts training. *Journal of Applied Developmental Psychology*, *25*(3), 283-302.

Leirhaug, P. E., & Annerstedt, C. (2016). Assessing with new eyes? Assessment for learning in Norwegian physical education. *Physical Education and Sport Pedagogy*, *21*(6), 616-631.

Leirhaug, P. E., & MacPhail, A. (2015). 'It's the other assessment that is the key': three Norwegian physical education teachers' engagement (or not) with assessment for learning. Sport, education and society, 20(5), 624-640.

Leirhaug, P. E., MacPhail, A., & Annerstedt, C. (2016). 'The grade alone provides no learning': investigating assessment literacy among Norwegian physical education teachers. Asia-Pacific Journal of Health, Sport and Physical Education, 7(1), 21-36.

Lewallen, T. C., Hunt, H., Potts-Datema, W., Zaza, S., & Giles, W. (2015). The whole school, whole community, whole child model: A new approach for improving educational attainment and healthy development for students. Journal of School Health, 85(11), 729-739.

Ling, J., Robbins, L. B., Wen, F., & Peng, W. (2015). Interventions to increase physical activity in children aged 2-5 years: a systematic review. Pediatric Exercise Science, 27(3), 314-333.

Livonen, S., Sääkslahti, A. K., Mehtälä, A., Villberg, J. J., Soini, A., & Poskiparta, M. (2016). Directly observed physical activity and fundamental motor skills in four-year-old children in day care. European Young children Education Research Journal, 24(3), 398-413.

Logan, S. W., Webster, E. K., Getchell, N., Pfeiffer, K. A., & Robinson, L. E. (2015). Relationship between fundamental motor skill competence and physical activity during childhood and adolescence: A systematic review. *Kinesiology Review*, *4*(4), 416-426.

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.

Longmuir, P. E., Gunnell, K. E., Barnes, J. D., Belanger, K., Leduc, G., Woodruff, S. J., & Tremblay, M. S. (2018). Canadian Assessment of Physical Literacy Second Edition: a streamlined assessment of the capacity for physical activity among children 8 to 12 years of age. BMC public health, 18(2), 1047.

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.

Lonsdale, C., Rosenkranz, R. R., Sanders, T., Peralta, L. R., Bennie, A., Jackson, B., ... & Lubans, D. R. (2013). A cluster randomized controlled trial of strategies to increase adolescents' physical activity and motivation in physical education: Results of the Motivating Active Learning in Physical Education (MALP) trial. *Preventive medicine*, *57*(5), 696-702.

Lopes, V. P., Barnett, L. M., Saraiva, L., Gonçalves, C., Bowe, S. J., Abbott, G., & Rodrigues, L. P. (2016). Validity and reliability of a pictorial instrument for assessing perceived motor competence in Portuguese children. Child: care, health and development, 42(5), 666-674.

López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2013). Alternative assessment in physical education: a review of international literature. *Sport, Education and Society*, *18*(1), 57-76.

Lounsbery, M. A., & McKenzie, T. L. (2015). Physically literate and physically educated: A rose by any other name?. *Journal of Sport and Health Science*, *4*(2), 139-144.

Lu, C., & Montague, B. (2016). Move to learn, learn to move: Prioritizing physical activity in young children education programming. *Young children Education Journal*, *44*(5), 409-417.

Lubans, D., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., ... & Biddle, S. (2016). Physical activity for cognitive and mental health in youth: a systematic review of mechanisms. *Pediatrics*, *138*(3), e20161642.

Lundvall, S. (2015). Physical literacy in the field of physical education–A challenge and a possibility. *Journal of Sport and Health Science*, *4*(2), 113-118.

Ma, X., Shen, J., Krenn, H. Y., Hu, S., & Yuan, J. (2016). A meta-analysis of the relationship between learning outcomes and parental involvement during young children education and early elementary education. Educational Psychology Review, 28(4), 771-801.

Mackintosh, K. A., Knowles, Z. R., Ridgers, N. D., & Fairclough, S. J. (2011). Using formative research to develop CHANGE!: a curriculum-based physical activity promoting intervention. *BMC public health*, *11*(1), 831.

Makopoulou, K., Neville, R. D., Ntoumanis, N., & Thomas, G. (2019). An investigation into the effects of short-course professional development on teachers' and teaching assistants' self-efficacy. Professional Development in Education, 1-16.

Mandigo, J., Francis, N., Lodewyk, K., & Lopez, R. (2012). Physical literacy for educators. Physical & Health Education Journal, 75, 27-30

Manios, Y., Moschandreas, J., Hatzis, C., & Kafatos, A. (1999). Evaluation of a health and nutrition education program in primary school children of Crete over a three-year period. *Preventive medicine*, *28*(2), 149-159.

Maude, P. (2010). Physical literacy and the young child. In Whitehead, M. (Ed.). (2010). *Physical literacy: Throughout the lifecourse* (pp. 100-115). London, UK: Routledge.

McEvoy, E., Heikinaro-Johansson, P., & MacPhail, A. (2017). Physical education teacher educators' views regarding the purpose (s) of school physical education. *Sport, Education and Society*, *22*(7), 812-824.

McWhannell, N., Triggs, C., & Moss, S. (2019). Perceptions and measurement of playtime physical activity in English primary school children: The influence of socioeconomic status. European Physical Education Review, 25(2), 438-455.

Merleau-Ponty, M. (1968). *The visible and the invisible: Followed by working notes*. Northwestern University Press.

Mitchell, J., Skouteris, H., McCabe, M., Ricciardelli, L. A., Milgrom, J., Baur, L. A., ... & Dwyer, G. (2012). Physical activity in young children: a systematic review of parental influences. Early Child Development and Care, 182(11), 1411-1437.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews, 4(1), 1.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & De Vet, H. C. (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.

Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Quality of Life Research, 27(5), 1171-1179.

Moreno, T. (2013). American physical education: A discursive essay on the potential unifying role of physical literacy in the United States. Journal of Sport Science and Physical Education, 65, 371-377.

Morgan, P., & Bourke, S. (2008). Non-specialist teachers' confidence to teach PE: the nature and influence of personal school experiences in PE. *Physical Education and Sport Pedagogy*, *13*(1), 1-29.

Morgan, P. J., & Hansen, V. (2008). Classroom teachers' perceptions of the impact of barriers to teaching physical education on the quality of physical education programs. Research quarterly for exercise and sport, 79(4), 506-516.

Morley, D., Van Rossum, T., Richardson, D., & Foweather, L. (2019). Expert recommendations for the design of a children's movement competence assessment tool for use by primary school teachers. *European Physical Education Review*, *25*(2), 524-543.

Murphy, C., Lundy, L., Emerson, L., & Kerr, K. (2013). Children's perceptions of primary science assessment in England and Wales. *British Educational Research Journal*, *39*(3), 585-606.

Myers, B. M., & Wells, N. M. (2015). Children's physical activity while gardening: development of a valid and reliable direct observation tool. Journal of Physical Activity and Health, 12(4), 522-528.

Ng, J. Y., Ntoumanis, N., Thøgersen-Ntoumani, C., Deci, E. L., Ryan, R. M., Duda, J. L., & Williams, G. C. (2012). Self-determination theory applied to health contexts: A meta-analysis. Perspectives on Psychological Science, 7(4), 325-340.

Ní Chróinín, D., & Cosgrave, C. (2013). Implementing formative assessment in primary physical education: teacher perspectives and experiences. *Physical Education and Sport Pedagogy*, *18*(2), 219-233.

Noonan, R. J., Boddy, L. M., Fairclough, S. J., & Knowles, Z. R. (2016). Write, draw, show, and tell: a child-centred dual methodology to explore perceptions of out-of-school physical activity. *BMC public health*, *16*(1), 326.

O'Loughlin, J., Chróinín, D. N., & O'Grady, D. (2013). Digital video: The impact on children's learning experiences in primary physical education. European Physical Education Review, 19(2), 165-182.

O'Sullivan, M. (2013). New directions, new questions: Relationships between curriculum, pedagogy, and assessment in physical education.

Owen, K. B., Smith, J., Lubans, D. R., Ng, J. Y., & Lonsdale, C. (2014). Self-determined motivation and physical activity in children and adolescents: A systematic review and meta-analysis. *Preventive medicine*, *67*, 270-279.

Paine, A. (2013). The effects of National Standards on the teaching of health and physical education in the New Zealand Curriculum. Masters of Education: University of Auckland.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74-98.

Parker, M., MacPhail, A., O'Sullivan, M., Ní Chróinín, D., & McEvoy, E. (2018). 'Drawing'conclusions: Irish primary school children's understanding of physical education and physical activity opportunities outside of school. European Physical Education Review, 24(4), 449-466.

Penney, D., Jones, A., Newhouse, P., & Cambell, A. (2012). Developing a digital assessment in senior secondary physical education. Physical Education and Sport Pedagogy, 17(4), 383-410.

Pérez, L. M. R., & Sanz, J. L. G. (2005). New measure of perceived motor competence for children ages 4 to 6 years. *Perceptual and motor skills*, *101*(1), 131-148.

Physical Literacy for Life (PLFL). (2017). Developing physical literacy and delivering quality sport. Retrieved from http://physicalliteracy.ca/

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J. P., Janssen, I., ... & Sampson, M. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition, and Metabolism*, *41*(6), S197-S239.

Popper, K. (1957). Philosophy of science. In C.A. Mace (Ed.), British philosophy in the mid-century. London, UK: George Allen and Unwin.

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.

Poucher, Z. A., Tamminen, K. A., Caron, J. G., & Sweet, S. N. (2019). Thinking through and designing qualitative research studies: a focused mapping review of 30 years of qualitative research in sport psychology. International Review of Sport and Exercise Psychology, 1-24.

Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research, 27(5), 1147-1157.

Public Health & Education (PHE) Canada. (2017). Physical literacy. PHE Canada. Retrieved from http://www.phecanada.ca/programs/physical-literacy

Quennerstedt, M. (2019). Physical education and the art of teaching: transformative learning and teaching in physical education and sports pedagogy. Sport, Education and Society, 24(6), 611-623.

Quennerstedt, M. (2019, June) Fantasmic logics of physical education. *Association Internationale des Écoles Supérieures d'Éducation Physique Conference,* Presentation given in Adelphi University, New York

Rainer, P., Griffiths, R., Cropley, B., & Jarvis, S. (2015). Barriers to Delivering Extracurricular School Sport and Physical Activity in Wales: A Qualitative Study of 5x60 Officers' Views and Perspectives. *Journal of Physical Activity and Health*, *12*(2), 245-252.

Renshaw, I., Chow, J.Y., Davids, K. and Hammond, J., 2010. A constraints-led perspective to understanding skill acquisition and game play: A basis for integration of motor learning theory and physical education praxis?. Physical Education and Sport Pedagogy, 15(2), pp.117-137.

Robertson, S., Kremer, P., Aisbett, B., Tran, J., & Cerin, E. (2017). Consensus on measurement properties and feasibility of performance tests for the exercise and sport sciences: a Delphi study. *Sports medicine-open*, *3*(1), 2.

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.

Robinson, D. B., Randall, L., & Barrett, J. (2018). Physical literacy (mis) understandings: What do leading physical education teachers know about physical literacy?. *Journal of Teaching in Physical Education*, *37*(3), 288-298.

Roetert, E. P., Kriellaars, D., Ellenbecker, T. S., & Richardson, C. (2017). Preparing students for a physically literate life. *Journal of Physical Education, Recreation & Dance*, *88*(1), 57-62.

Roetert, E. P., Ellenbecker, T. S., & Kriellaars, D. (2018). Physical literacy: why should we embrace this construct?.

Rooney, L. E., Videto, D. M., & Birch, D. A. (2015). Using the whole school, whole community, whole child model: implications for practice. *Journal of school health*, *85*(11), 817-823.

Rosa, R. L., Ridgers, N. D., & Barnett, L. M. (2013). Development and use of an observation tool for active gaming and movement (OTAGM) to measure children's movement skill components during active video game play. Perceptual and motor skills, 117(3), 935-949.

Roscoe, C. M., James, R. S., & Duncan, M. J. (2019). Accelerometer-based physical activity levels, fundamental movement skills and weight status in British preschool children from a deprived area. *European journal of pediatrics*, 1-10.

Rudd, J., Butson, M. L., Barnett, L., Farrow, D., Berry, J., Borkoles, E., & Polman, R. (2016). A holistic measurement model of movement competency in children. Journal of Sports Sciences, 34(5), 477-485.

Santos-Beneit, G., Sotos-Prieto, M., Bodega, P., Rodríguez, C., Orrit, X., Pérez-Escoda, N., ... & Penalvo, J. L. (2015). Development and validation of a questionnaire to evaluate lifestyle-related behaviors in elementary school children. *BMC public health*, *15*(1), 901.

Schools and Physical Activity Task and Finish Group. (2013). Physical literacy – An all-Wales approach to increasing levels of physical activity for children and young people. Cardiff, UK: Crown.

SHAPE America. (2014). National standards & grade-level outcomes for K-12 physical education. Champaign, IL: Human Kinetics

SHAPE America. (2017). Physical literacy. Retrieved from http://www.shapeamerica.org/events/physicalliteracy.cfm.

Shepard, L. A. (2000). The role of assessment in a learning culture. Educational researcher, 29(7), 4-14.

Shewmake, C. J., Merrie, M. D., & Calleja, P. (2015). Xbox kinect gaming systems as a supplemental tool within a physical education setting: Third and fourth grade students' perspectives. The Physical Educator, 72(5).

Smith, B., & McGannon, K. R. (2018). Developing rigor in qualitative research: Problems
    and opportunities within sport and exercise psychology. International review of sport
    and exercise psychology, 11(1), 101-121.

Speyer, R., Kim, J. H., Doma, K., Chen, Y. W., Denman, D., Phyland, D., ... & Cordier, R.
    (2019). Measurement properties of self-report questionnaires on health-related quality
    of life and functional health status in dysphonia: a systematic review using the
    COSMIN taxonomy. Quality of Life Research, 28(2), 283-296.

Sport England. (2016). Sport England: Towards an active nation. Retrieved from
    https://www.sportengland.org/media/10629/sport-england-towards-an-active-
    nation.pdf

Sport England. (2019). Active lives children and young people survey: Attitudes towards
    sport and physical activity. Retrieved from
    https://www.sportengland.org/media/13851/active-lives-children-survey-2017-18-
    attitudes-report.pdf

Sport New Zealand (2015). Physical literacy approach - Guidance for quality physical
    activity and sport experiences. Retrieved from
    http://www.sportnz.org.nz/assets/Uploads/attachments/About-us/2015-
    PhysicalLiteracyDocument-Online.pdf

Sport New Zealand (2016). Annual report 2016. Retrieved from
    http://www.sportnz.org.nz/assets/Uploads/3019-SNZ-Annual-Report-2016-FINAL-
    singlepages2.pdf

Sport Wales. (2017). Physical literacy programme for schools (PLPS) | community sport |
    sport Wales - Chwaraeon Cymru. Retrieved from http://sport.wales/community-
    sport/education/physical-literacy-programme-for-schools-(plps).aspx

Sprake, A., & Walker, S. (2013). "Strike While the Iron is Hot": the duty of physical
    education to capitalise on its' compulsory position with a holistic curriculum
    underpinned by physical literacy. *Presidents Message 12*.

Sturgess, J., & Ziviani, J. (1996). A self-report play skills questionnaire: Technical development. *Australian Occupational Therapy Journal*, *43*(3-4), 142-154.

Sun, H. (2015). Operationalizing physical literacy: the potential of active video games. Journal of Sport and Health Science, 4(2), 145-149.

Taplin, L. (2013). Physical literacy as journey. *ICSSPE Journal of Sport Science and Physical Education, Bulletin*, *65*, 57-63.

Telama, R., Yang, X., Viikari, J., Välimäki, I., Wanne, O., & Raitakari, O. (2005). Physical activity from childhood to adulthood: a 21-year tracking study. *American journal of preventive medicine*, *28*(3), 267-273.

Tenenbaum G, Eklund RC, Kamata A. Measurement in sport and exercise psychology. Human Kinetics; 2018 Nov 15

Terwee, C. B., Prinsen, C. A., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., ... & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Quality of Life Research, 27(5), 1159-1170.

The Aspen Institute. (2013). Project play. Retrieved from https://www.aspenprojectplay.org/

The Aspen Institute. (2015). Physical literacy in the United States: A model, strategic plan, and call to action. Retrieved from http://aspenprojectplay.org/sites/default/files/PhysicalLiteracy_AspenInstitute.pdf

The Sport for Life Society. (2017). Physical Literacy. Retrieved from http://sportforlife.ca/

Tinner, L., Kipping, R., White, J., Jago, R., Metcalfe, C., & Hollingworth, W. (2019). Cross-sectional analysis of physical activity in 2–4-year-olds in England with paediatric quality of life and family expenditure on physical activity. *BMC public health*, *19*(1), 846.

Tolgfors, B. (2018). Different versions of assessment for learning in the subject of physical education. *Physical Education and Sport Pedagogy*, *23*(3), 311-327.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International journal for quality in health care*, *19*(6), 349-357.

Torrance, H., & Pryor, J. (1998). Investigating formative assessment: Teaching, learning and assessment in the classroom. McGraw-Hill Education (UK).

Tremblay, M. S., Costas-Bradstreet, C., Barnes, J. D., Bartlett, B., Dampier, D., Lalonde, C., ... & Way, R. (2018). Canada's Physical Literacy Consensus Statement: process and outcome. *BMC Public Health*, *18*(2), 1034.

Tremblay, M., & Lloyd, M. (2010). Physical literacy measurement-the missing piece. *Physical and health education journal*, *76*(1), 26-30.

Tremblay, M. S., & Longmuir, P. E. (2017). Conceptual critique of Canada's physical literacy assessment instruments also misses the mark. *Measurement in Physical Education and Exercise Science*, *21*(3), 174-176.

Tsangaridou, N. (2016). Moving towards effective physical education teacher education for generalist primary teachers: a view from Cyprus. *Education 3-13*, *44*(6), 632-647.

UKactive., (2018) Generation inactive 2. Retrieved from https://www.ukactive.com/wp-content/uploads/2018/09/Generation_Inactive-2_Nothing_About_Us_Without_Us.pdf

Ulrich, D. A. (2013). The test of gross motor development-3 (TGMD-3): Administration, scoring, and international norms. Spor Bilimleri Dergisi, 24(2), 27-33.

Vandorpe, B., Vandendriessche, J., Lefèvre, J., Pion, J., Vaeyens, R., Matthys, S., ... & Lenoir, M. (2011). The Körperkoordinationstest für kinder: Reference values and suitability for 6–12-year-old children in Flanders. *Scandinavian journal of medicine & science in sports*, *21*(3), 378-388.

van Rossum, T., Foweather, L., Richardson, D., Hayes, S. J., & Morley, D. (2019). Primary
     Teachers' Recommendations for the Development of a Teacher-Oriented Movement
     Assessment Tool for 4–7 Years Children. *Measurement in Physical Education and
     Exercise Science*, *23*(2), 124-134.

van Rossum, T., & Morley, D. (2018). The role of digital technology in the assessment of
     children's movement competence during primary school physical education
     lessons. *Digital Technology in Physical Education: Global Perspectives*.

Wainwright, N., Goodway, J., Whitehead, M., Williams, A., & Kirk, D. (2018). Laying the
     foundations for physical literacy in Wales: the contribution of the Foundation Phase to
     the development of physical literacy. *Physical Education and Sport Pedagogy*, *23*(4),
     431-444.

Warburton, D. E., & Bredin, S. S. (2017). Health benefits of physical activity: a systematic
     review of current systematic reviews. *Current opinion in cardiology*, *32*(5), 541-556.

Weikart, P. S., Schweinhart, L. J., & Larner, M. (1987). Movement curriculum improves
     children's rhythmic competence. HighScope ReSource, 6(1), 8-10.

Weinburg, B. (Ed.) (2013). Feature: Physical literacy (Special issue). Journal of Sport
     Science and Physical Education, 65, 18-20.

World Health Organization. (2019). *Global action plan on physical activity 2018-2030: more
     active people for a healthier world*. World Health Organization.

Whitehead, M. (1993, August). Physical literacy. Paper presented at the IAPESWG
     Congress, Melbourne, Australia.

Whitehead, M. (2001). The concept of physical literacy. European Journal of Physical
     Education, 6, 127-138.

Whitehead, M. (2007). Physical literacy: Philosophical considerations in relation to developing a sense of self, universality and propositional knowledge. Sports, Ethics, and Philosophy, 1, 281-298.

Whitehead, M. (Ed.). (2010). *Physical literacy: Throughout the lifecourse.* London, UK: Routledge.

Whitehead, M. (2013a). The history and development of physical literacy. ICSSPE Journal of Sport Science and Physical Education, 65, 22-28.

Whitehead, M. (2013b). Definition of physical literacy and clarification of related issues. ICSSPE Journal of Sport Science and Physical Education, 65, 29-34. Whitehead 1993

Whitehead, M. (2013). ICSSPE Bulletin, 65(1.2).

Whitehead, M. (2019). Physical literacy across the world. London, UK: Routledge.

Wiliam, D., & Thompson, M. (2017). Integrating assessment with learning: What will it take to make it work?. In The future of assessment (pp. 53-82). Routledge.

Wiseman, N., Harris, N., & Downes, M. (2017). Validation of an iPad activity to measure preschool children's food and physical activity knowledge and preferences. *international journal of behavioral nutrition and physical activity*, *14*(1), 11.

Yao, C. A., & Rhodes, R. E. (2015). Parental correlates in child and adolescent physical activity: a meta-analysis. International journal of behavioral nutrition and physical activity, 12(1), 10.

Youth Sport Trust (2013). Primary physical literacy framework. Retrieved from https://www.youthsporttrust.org/sites/yst/files/resources/documents/physical_literacy_framework.pdf

Zetou, E., Nikolaos, V., & Evaggelos, B. (2014). The effect of instructional self-talk on performance and learning the backstroke of young swimmers and on the perceived functions of it. Journal of Physical Education and Sport, 14(1), 27.

Appendix One

PICO information and search strand

for the systematic review (study 2)

PICO (Population Intervention Context Ouctomes) information

	Include	Exclude
Population	Typically developing children Age 3-11	Is not used with children Special Populations
Intervention	Studies will be included if they report an Assessment OR Measurement OR Test OR Tool OR Instrument OR Battery OR Method OR Psychometric OR Observation OR Indicator OR Evaluation OR Validity Or Reliability	
Context	Physical literacy, physical activity, play, sport, physical education, exercise, recreation Field Based assessment	Lab based
Outcomes	Assessment of outcome(s) related to physical literacy. Motivation OR Enjoyment OR Confidence OR Self Or "Perceived Competence" OR Affective OR Social OR Emotion OR Attitude OR Belief OR Physical OR Fitness OR Motor OR Movement OR Skills OR Technique OR Mastery OR Ability OR Coordination OR Performance OR "Perceptual Motor" OR Knowledge OR Understanding OR Value OR Cognitive OR Health OR Wellbeing	
Study design/publication characteristics	Peer reviewed journal articles published in English	Studies published in a foreign language Not published in a peer reviewed journal Duplicate publication They did not assess the psychometric properties of the relevant instrument Full text articles were not available, Studies that dealt with translated and culturally adapted versions of the measures Book chapters, case studies, student dissertations, conference abstracts, review articles, meta-analyses and editorials , protocol papers, systematic reviews Named and used in multiple studies

Search strand for the systematic review

	Assess* OR Measure* OR Test* OR Tool* OR Instrument* OR Battery* OR Method* OR Psychometr* OR Observ* OR Indicator* OR Evaluat* OR Valid* Or Reliab*	Title or Abstract
AND	"Physical* Activ*" OR "Physical* Liter*" OR Play OR Sport OR "Physical* Educat*" OR Exerci* OR Recreation	All Text
AND	Child* OR Youth OR Adoles* OR P$ediatric* OR Schoolchild* OR Boy* OR Girl* OR Preschool* OR Juvenile* OR Teenager	All Text
AND	Motiv* OR Enjoy* OR Confidence OR Self* Or "Perc* Competence" OR Affective OR Social OR Emotion* OR Attitude* OR Belief* OR Physical* OR Fitness OR Motor OR Movement* OR Skills* OR Technique* OR Mastery OR Abilt* OR Coordination OR Performance OR "Perceptual Motor" OR Knowledge OR Understanding OR Value OR Cogniti* OR Health OR Well*	All Text

Appendix Two

Focus Group Guides

Teachers and Expert Focus Group Guides

In the following questions, I would like to understand your own opinions of assessment and physical literacy, please be as honest as possible

1. Is it important to assess physical literacy? **5 minutes**
    Prompts:
    a. Is it important to assess to affective/cognitive/physical?
    b. Who would find this information useful? Why?
    c. What could this information be useful for?

    So we've just talked a bit about your views in the importance of physical literacy assessment, I now want to focus on the purpose and function of an assessment

2. How can the assessment of physical literacy best support pedagogy, teaching and learning? **5 minutes**

    So we've just talked a bit about your views in why we should do it, but this doesn't always translate to what we're actually able to do in practice. So I'm now going to ask a little bit about that

3. Do you have any experience of assessing physical literacy? **10 minutes**
    a. Can you explain what an assessment you may have used looks like? (If not offered-pushed for what is assessed and how)
    b. Are you aware of any physical literacy assessments?
    c. Who conducts the assessment? How long does it take? Does it need much equipment? How often would you use this assessment? Level of expertise required? What did the coaches/teachers/children think of the assessment?
    d. How is the assessment scored? How is the information fedback and to whom?
    e. Who uses this information? Why?
    f. What are the best and worst aspects of these assessments, and why?

    We know that there are many barriers to assessment to physical literacy in schools. These include time, lack of importance, space, lack of resources, lack of confidence. Rather than focussing on the barriers we know are there, The next question is focussing on solutions

4. **Solutions task (See attached):** How can we overcome the barriers around physical literacy assessment? **10 minutes**

So we've spoken about your positive and negative experiences of assessment, and how we may overcome some of the existing barriers. Now we'd like to start talking about the development of a new assessment of physical literacy for use with children

Tick sheet task (see attached) 10 minutes: *As part of our systematic review, we have identified a number of 'sub-elements' within each domain, that currently appear in existing assessments. In the table below, please rank your perception of the relative importance of these 'sub-elements' for both children aged 3-7 years old, and children aged 7-11 years old.*

45-55 minutes

Design task (see attached) 15 minutes-What do you think the ideal physical literacy assessment would look like?

To be introduced by HG and CS, but led my moderators on tables

Taking on those really valid positive and negative aspects of current assessments, In the next questions, I'd like to talk about an ideal assessment.

Please use the pens and flip chart paper to design this. After 10 minutes, each group will have 1 minute to present their ideas back to the group.

a. What assessment approaches (Includes techniques, tools, strategies) could be used to effectively assess physical literacy?
b. What could this look like in practice?
c. Who would be leading **and why**? (teacher/peer/self/others)
d. How would it be administered and conducted?
e. What support/materials would the children need for this?
f. How long would it take?
g. How often would you do it?
h. How would you track information?
i. Feedback and results – how can they be easily accessed and understood by users?
j. How can technology be used to support this assessment of physical literacy?
k. What would be the strengths of this method?
l. Can you think of any limitations to this?

HG and CS to give warning at 8.30

55-60 minutes: HG and CS to facilitate group feedback and Dot-mocracy

Child Focus Group Guides

So I can try and learn some of your names, and know who you are on the tape, can you please tell me your name and then tell me your favourite game?
OK so today we are going to talk a bit about games, PE and all different types of activity. First, I'd like to talk about all different ways you can tell if you are good at something.

1. **Drawing Task:** *Can you draw me a picture about a time you were assessed or tested in PE?*
   How do you know you are doing well?
   What would it look like if you were not doing very well?
   How often do you do something like this in PE?
   Who else is in the picture?
   Is there any way teachers check how good you are in PE?

What are the best parts about a test? Why?
What are the worst parts about a test? Why?
   Can you give me an example of how a teacher would tell you have done well in PE?
   Do teachers ever check if you've had fun?
   Do teachers ever check if you understand why you're doing PE?
   What about in other types of games?

So we've just talked about the ways a teacher might be able to tell if you've done well or not so well in PE. I'd like you to keep thinking about these

2. What parts of these tests did you like/not like?
   What would make a test better? Why?
   What are the worst parts of a test? Why?
   What would make a test better? Why?

3. **Picture Task:** *So now we'd like your help! We are going to make a test to help us find out all the ways children like you think, feel and about physical activity, and what you can do.*
Physical activity is any sort of movement we do that burns energy. It can be games, play, PE, sport, walking, cycling, swimming, and lots of other things that are fun and keep us healthy. We are going to have a look at some pictures/video to help us think about this.

   What test could we make that would show us these things?
   If you were watching from the side what would be going on?
   Who would be in charge?
   What equipment would you need?
   How long would it take?
   How often would you do it?
   How would you keep track of all the information?
   Who would find that information useful?

We've spoken a lot about how things you think, feel and do to help you be active and how you might show or tell a grown up these things. Now I want you to think about you on your own being active. Remember, it can be games, play, PE, sport, walking, cycling, swimming, and lots of other things.

4. How would you tell if you've done well being active without a teacher/adult to help?
   Do you ever think about how you've done in PE?
   How does it make you feel?
   Can this help you get better?

*There are things we **think** that helps us be active.*

These people knows why physical activity is good for them

This person knows what physical activity does to the body

These people know lots of different ways to be active

*There are things we **feel** that help us be active.*

This person is brave when playing games

This person really wants to be active

These people are happy when they are active   This person keeps going when the find something hard

*There are things we **do** that help us be active.*

This person is good at balancing

These people can play with balls in lots of different ways

These people can move in lots of different ways

Appendix 3

Personal characteristics of contributors

Dr Lawrence Foweather	Senior Lecturer in Physical Activity, Exercise and Health in the School of Sport and Exercise Sciences. He is a member of the Physical Activity Exchange within the Research Institute for Sport and Exercise Sciences.

- ❖ 2014, Liverpool John Moores University, UK, Postgraduate Certificate Teaching and Learning in Higher Education
- ❖ 2010, Liverpool John Moores University, UK, PhD in Fundamental Movement Skills, Physical Activity and Psychological Well-being Among 9-10 Year old Children
- ❖ 2005, Loughborough University, UK, MSc Physical Activity and Health
- ❖ 2002, De Montfort University, UK, BSc Sports Studies

Dr Lynne M Boddy	Reader in Children's Physical Activity and Physical Activity Exchange Lead based within the Research Institute for Sport and Exercise Sciences.

- ❖ 2009, Liverpool John Moores University, United Kingdom, PhD: An investigation into the body composition and physical fitness of Liverpool Schoolchildren, Liverpool John Moores University, 2009.
- ❖ 2009, Liverpool John Moores University, United Kingdom, PGCert Learning and Teaching in Higher Education (Distinction)
- ❖ 2002, Liverpool John Moores University, United Kingdom, BSc(hons) Sports Science (Physiology), 2.1

Professor Zoe Knowles	Professor Zoe Knowles is a HCPC Registered Practitioner Psychologist; Associate Fellow of the British Psychological Society (AFBPsS) and Fellow of the British Association of Sport and Exercise Sciences (FBASES). Psychology Division Chair for BASES and Board member. Strategic Lead for External Engagement in the School of Sport and Exercise Sciences at Liverpool John Moores University.

- ❖ Professor of Engagement and Learning, Sport and Exercise Sciences, Liverpool John Moores University, 2017 – present
- ❖ Reader of Sport and Exercise Psychology, Physical Activity Exchange at the School of Sport and Exercise Sciences, Liverpool John Moores University, 2013 – present
- ❖ L/SL in Sport and Exercise Psychology, School of Sport and Exercsie Sciences, Liverpool John Moores University, 1998 - 2013

Dr Paula Watson Senior lecturer in exercise and health psychology, and a member of the Physical Activity Exchange in the Research Institute for Sport and Exercise Sciences. Chartered psychologist with the British Psychological Society and registered as a sport and exercise psychologist with the Health and Care Professions Council.

- ❖ 2014, Liverpool John Moores University, United Kingdom, Postgraduate Certificate in Learning and Teaching in Higher Education (Distinction)
- ❖ 2012, Liverpool John Moores University, United Kingdom, PhD: Feasibility evaluation and long-term follow up of a family-based behaviour change intervention for overweight children (GOALS)
- ❖ 2004, Liverpool John Moores University, United Kingdom, MSc Sports Psychology (Distinction)
- ❖ 2000, University of York, United Kingdom, BSc (Hons) Psychology (1st class)