CRediT authorship contribution statement

Katie Fitton Davies: Conceptualisation, Methodology, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Validation. **Paula M Watson:** Conceptualisation, Methodology, Validation, Writing – review & editing. **James R Rudd:** Conceptualisation, Methodology, Project administration, Supervision, Writing – review & editing. **Simon Roberts:** Conceptualisation, Methodology, Supervision, Validation, Writing – review & editing. **Farid Bardid:** Conceptualisation, Methodology, Writing – review & editing. **Zoe Knowles:** Conceptualisation, Methodology, Writing – review & editing. **Lawrence Foweather:** Conceptualisation, Methodology, Project administration, Supervision, Writing – review & editing.

Development and validity of the Motivation Assessment Tool for Physical Education (MAT-PE) among young children

Authors: Fitton Davies, K.[1,2], Watson, P. M.[2], Rudd, J. R.[1], Roberts, S.[1], Bardid, F.[3,4], Knowles, Z.[2], & Foweather, L.[2]

Affiliations:

1. Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, IM Marsh, Barkhill Road, Liverpool, United Kingdom L176BD
2. Physical Activity Exchange, Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, 5 Primrose Hill, Liverpool, United Kingdom, L32EX
3. School of Education, University of Strathclyde, Glasgow, United Kingdom
4. Department of Movement and Sports Sciences, Ghent University, Ghent, Belgium

Corresponding author

Lawrence Foweather: l.foweather@ljmu.ac.uk; Physical Activity Exchange, Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, 5 Primrose Hill, Liverpool, United Kingdom, L32EX

Declaration of interest: none

Funding statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Introduction: It is important to understand young children's motivation within Physical Education (PE) so that researchers and teachers can effectively support children's physical, affective, social and cognitive development as well as physical activity (PA) behaviors. However, there is a dearth of motivation research in PE with children under the age of seven due to a lack of developmentally appropriate assessment tools. **Aims:** This multi-study paper outlines the development content and construct validity of a novel, mixed-method tool to assess young children's psychological needs and behavioral regulation within PE (Motivation Assessment Tool for Physical Education; MAT-PE). **Methods:** Study 1 consisted of the iterative development of the MAT-PE through working with 43 young children (ages 5-6) from three primary schools located within a large city in North West England. This work culminated in MAT-PE version 1, which was examined for content validity in a further sample of 85 children (ages 5-6) from 12 primary schools located within a large city in North West England. Study 2 consisted of the development, content validation, acceptability and inter- and intra-rater reliability of the MAT-PE codebook. Study 3 explored construct validity through hypothesis-testing via correlational data. Descriptive data captured through the MAT-PE and codebook with 78 children (ages 5-6) from 12 primary schools located within a large city in North West England is also presented. **Findings:** The MAT-PE and its codebook were judged to have promising content validity, the codebook was deemed acceptable, as well as demonstrating excellent inter- and intra-rater reliability (ICC = .90). Regarding construct validity, as hypothesised, all psychological needs were positively correlated and autonomous regulations were negatively associated with amotivation. There were also unexpected correlations such as the negative correlation between intrinsic and identified regulation. **Conclusion:** Further development of the MAT-PE is required; however, this study has taken a promising first step in developing a tool to comprehensively measure five- to six-year-old children's motivational perceptions in PE.

Keywords: self-determination theory; physical education; children; mixed methods; codebook, assessment

Introduction

Physical Education (PE) supports physical, affective, social and cognitive development for primary school aged children (5-11-years-old) and promotes healthy lifestyles (Bailey, 2006; Casey & Goodyear, 2015; Hills et al., 2015; Loprinzi et al., 2015; Marques et al., 2017; Sallis et al., 1991, 2012; Tsangaridou & Lefteratos, 2013). Focusing on the affective domain, PE provides a context to foster children's perceived competence, motivation and enjoyment in physical activity (PA) and movement (Carroll & Loumidis, 2001; Chen, 2014). Early learning experiences in PE are thus considered critical for continued participation in PA (Hills et al., 2015; Kirk, 2005), with enjoyment of PE positively affecting future attitudes and intention towards PA (Ladwig et al., 2018). Motivation is a mechanism that helps sustain behaviour and engagement within PE. Therefore, understanding how to foster and maintain motivation in children within primary PE is key for supporting their PA participation (Jaakkola et al., 2013; Standage et al., 2003), physical literacy and well-being (Whitehead, 2019).

Guay et al. (2010) demonstrated that children aged six to ten years report on their motivation differently between school subjects, highlighting the importance of assessing children's motivation according to specific subjects. Despite variances in cognitive ability and communication skills, young children (aged 4-7 years) are able to recognise the subject of PE as a forum for learning how to move their bodies, to exercise and get fit, and can recall activities completed during PE lessons (Solmon & Carter, 1995). As such, the present paper is concerned with young children's contextual motivation toward PE (Vallerand, 1997, 2007). Specifically, we were interested in young children's ability to conceptualise a) the motivating factors driving their PE behaviours, and b) the social-contextual factors within the PE environment that relate to the satisfaction of autonomy, competence and relatedness. These conceptualisations are the central tenets of Organismic Integration Theory (OIT) and Basic Psychological Needs Theory (BPNT) respectively, which of the six mini-theories within Self-Determination Theory (SDT; Ryan & Deci, 2017), are arguably the most widely used in PA (Teixeira et al., 2012) and PE research (Vasconcellos et al., 2019).

OIT focuses upon internalisation and integration, resulting in different types of motivation that vary in their degree of autonomy as well as in their specific antecedents and effects on experience and behaviour within a socio-cultural environment such as PE. OIT is centred around the tenet that some behavioral regulations are experienced as "relatively alien to the self" while others are more "autonomously enforced" (Ryan & Patrick 2009, p. 112), whereby extrinsic motivation lies upon a continuum of autonomy. After amotivation (no motivation to act) are two forms of controlled motivation characterised by pressured engagement in an activity: external regulation (driven by reward or avoidance of punishment and considered the least internalised form of motivation), and introjected regulation (driven by the ego/pride or guilt/shame). Following with increasing degrees of internalisation are identified regulation (driven by a desire to pursue an internal goal) and integrated regulation (driven by aligned values and behaviours). Together with intrinsic regulation (driven by inherent pleasure, interest or challenge), identified and integrated regulation are forms of autonomous motivation, characterised by levels of volition and self-endorsement (Ryan & Deci, 2017).

For children to flourish in wellbeing and performance, three basic psychological needs (BPN) must be supported and satisfied within the social environment, leading to autonomous motivation (Katz et al., 2011; Milyavskaya & Koestner, 2011; Standage et al., 2012). The needs are *competence* (the need for satisfaction in demonstrating capabilities), *autonomy* (the need for actions to be volitional and a sense of choicefulness (Vansteenkiste et al., 2005)) and *relatedness* (the need to seek out connected relationships with others: Deci & Ryan, 2000). Past research in older children have shown that children perceive higher levels of relatedness and more moderate levels of autonomy (Ntoumanis et al., 2009; van Aart et al., 2017), while it is common to find higher competence levels in younger (Barnett et al., 2015; Spessato et al., 2013). Thus, the extent of internalisation (and the quality of motivation) and need satisfaction experienced by a child in PE is dependent upon the extent to which the three BPN are supported by their PE teacher's delivery style and the PE environment. Autonomy can be supported by providing meaningful choices,

competence by providing guidance, and relatedness by providing a friendly demeanour (Ryan & Deci, 2017). Thus, autonomy, competence and relatedness act as mediators between the contextual factors (PE teacher and children's peers) and contextual motivation (intrinsic, extrinsic and amotivation) (Vallerand, 1997, 2007).

Across the globe, research supports the use of SDT as a framework for supporting positive experiences and participation in PE. In the USA, Erwin et al. (2013) found that autonomy support (choice vs no choice) and lesson structure (individual vs group activities) affected PA levels during PE among 8-11-year-olds. Leptokaridou et al. (2016) found positive relationships between autonomy-supportive teaching and effort and enjoyment in PE among 10-12-year-olds from Greece, while Escriva-Boulley et al. (2018) reported a positive association between autonomy support and moderate-to-vigorous physical activity (MVPA) during PE in 5-11-year-olds from France. Within the UK, numerous studies have explored SDT in PE among youth (aged 11 to 16 years: Ntoumanis, 2005; Standage et al., 2003, 2005; Taylor & Ntoumanis, 2007). These studies also demonstrate that a need supportive motivating teaching style in PE leads to greater need satisfaction among students, which in turn predicts intrinsic motivation and future participation in PA inside (optional PE) and outside of school (leisure PA). However, to our knowledge, no study has explored young children's (5-7-year-olds) motivation for early primary school PE. This age period is important to understand, motivationally, as MVPA levels begin to decline from the age of school entry (Reilly, 2016). Furthermore, while previous literature in 8- to 12-year-olds has reported that motivation for PE, assessed using a 33 item Likert scale survey, declines with age (Chanal et al., 2019), it is important to understand whether this decrease occurs at an earlier age to put in place preventative actions. Given that children can differentiate between behavioral regulations far earlier than first posited (Butler, 2005), examining 5-7-year-olds motivation for PE warrants further study in order to investigate how different learning environments, motivational climates and PE teaching styles affect self-determined motivation through their impact on perceptions of competence, autonomy, and relatedness.

One of the reasons for a lack of research into young children's motivation is the paucity of measurement tools available for this younger age group (Sebire et al., 2013). Indeed, few tools exist specifically for use with young children. For instance, Gottfried (1990) adapted The Children's Academic Intrinsic Motivation Inventory (CAIMI; Gottfried, 1986) for use in younger children (ages 7-9; Gottfried, 1990). In another example, Guay et al. (2010) modified the Academic Motivation Scale (AMS; Vallerand et al., 1989) to create the Elementary School Motivation Scale (ESSMS) designed for use in 6-9-year-old children. However, it should be noted that these quantitative tools focused exclusively on intrinsic motivation (Gottfried, 1986, 1990), collapsed motivational constructs (Guay et al., 2010), omitted amotivation and were not PE specific. By isolating single components and grouping constructs into broader categories, these measures are insensitive to motivational intricacies and fail to provide a comprehensive assessment of young children's motivation. Furthermore, these surveys typically capture responses using Likert scales (except for the ESSMS which used a double-binary response system), which have been found to be unreliable among young children due to their limited cognitive understanding (Mellor & Moore, 2014). Gelman and Baillargeon (1983) argued that young children think dichotomously; thus, future research should incorporate alternative response formats into assessments (Mellor & Moore, 2014). Research exploring young children's perceived competence has demonstrated success in using structured alternative response formats and utilising pictures within their measurement tools (Harter & Pike, 1984; Barnett et al., 2015). Such research instruments could inform the design of assessments of motivation for PE within this age group.

Children as young as five years of age have been found to be able to describe their internal mental states such as their perceptions, emotions, cognitions and physiological states (Stone & Lemanek, 1990). This suggests that qualitative methodologies could be used to elicit young children's voices concerning 'why questions' for motivation in PE. Previous research (Chandler & Connell, 1987) has used a structured interview procedure and content analysis to explore behavioral regulations towards general 'liked' (e.g., playing a board game) and 'disliked' (e.g., going to bed on

time) behaviours amongst children aged 5-13-years-old. Importantly, this research showed that intrinsic, extrinsic and internalised forms of motivation are conceptually and developmentally distinct, and therefore should be explored separately within children's motivational research (not collapsed or omitted). However, while the methodology shows some promise, the study did not examine PE, amotivation was not considered, and the types of behavioral regulation were not clearly delineated. Other research has examined motivation for reading in 6-8-year-old children through qualitative case studies (Erickson, 2019), however, again, the study did not examine PE and the sample size was small due to the methodology (n=8). Qualitative methods published in other fields of research could offer promising approaches to assessing young children's motivation. For example, the Write and Draw technique alongside semi-structured interviews has been effectively used to capture views on passive smoking in children aged four to eight (Porcellato et al., 2005; Woods et al., 2005). Evolving this methodology, Noonan et al. (2016) developed a humanistic, child-led interactive method called the Write, Draw, Show and Tell which successfully gathered 10 to 11- year-old children's perspectives on PA and may offer a viable means by which to explore BPN and behavioral regulation in younger children. Developing a tool that can assess young children's motivation within PE would benefit researchers as it would improve understanding of the psychological mediators that affect young children's motivation and related contextual cognitive, affective and behavioral outcomes (Ferrer-Caja & Weiss, 2000) and as such inform intervention design. Furthermore, educational curricula aim to be more child-centred (Department of Education, 2014) but no appropriate tools for affective outcomes exist. A novel tool is therefore needed to better understand young children's motivation within PE which could inform teaching styles, bridging the gap between research and practice.

In summary, supporting children's motivation within PE is crucial for their holistic development (Bailey, 2006; Casey & Goodyear, 2015). Little is known about young children's (five- to six-year-old) motivation towards PE due to a lack of empirical studies (Vasconcelloset al., 2019), which is likely due to a lack of developmentally-appropriate measures (Sebire et al, 2013). To date,

quantitative and qualitative methods have been utilised separately in order to measure motivation, primarily within OIT, in academic subjects, and with older children. A mixed-method approach would provide a more comprehensive overview of motivation in PE among young children (Caruth, 2013). Therefore, we aimed to develop an age-appropriate, mixed-method tool aligned with SDT in order to measure young children's motivation in PE (Motivation Assessment Tool for Physical Education; MAT-PE).

This paper reports the initial development and content validity of the MAT-PE and its associated codebook. According to the Consensus-based Standards for the Selection of Health Instruments (COSMIN: Mokkink et al., 2010; Terwee et al., 2018), content validity, defined as 'the degree to which the content of an instrument is an adequate reflection of the construct to be measured' (Mokkink et al., 2010), is the most important measurement property of a tool and the key focus for tool development (Terwee et al., 2018). We also present preliminary descriptives to illustrate the MAT-PE data and an initial exploration of construct (structural) validity, another important measurement property for evaluating outcome measures (Prinsen et al., 2016). COSMIN guidelines state the need for *a priori* hypotheses for construct validity (Mokkink et al., 2012). Thus, based on SDT research (Ryan & Deci, 2017; Sebire et al., 2013; van Aert et al., 2017), broadly we hypothesised and expected that 1) BPN will positively associate with each other, 2) BPN will positively associate with autonomous regulations and negatively associate with controlled regulations and amotivation, 3) introjection will demonstrate complex associations with the other variables, and 4) behavioral regulations will ascribe to the simplex model (Ryan & Connell, 1989). This research is reported across three studies (Table 1). All studies took place within a wider cluster randomized controlled trial (RCT; Rudd et al., 2020) and were approved by the university research ethics committee (Ref. 17/SPS/031).

Table 1

Indicative content of the studies presented in this manuscript

Study 1	**Study 2**	**Study 3**
Development & content validity	*Analysis and scoring*	*Construct validity*

• Development of the MAT-PE	• MAT-PE Codebook development	• MAT-PE descriptive data
• Description of the MAT-PE	• Content validity and acceptability of the MAT-PE codebook	• Hypothesis-testing (correlations)
• Content validity of the MAT-PE	• Reliability of the MAT-PE codebook	

Note. MAT-PE=Motivation Assessment Tool for Physical Education

Study 1: Development and content validity of the MAT-PE

Method

Tool development

Supplementary Material A provides a detailed overview of the iterative development of the MAT-PE tool and resources. Briefly, methodological development was guided COSMIN, more specifically, COSMIN guidelines on content validity, which is a methodology developed via Delphi study including 159 experts from 21 countries in order to produce guidelines on content validity (COSMIN; Terwee et al., 2018). In accordance with this guidance and that of Dunn et al. (1999), a team of cross-disciplinary researchers (KFD, PW, JR, SR, FB, ZK, LF) constituted of Professors, Readers and Senior Lecturers with primary areas of expertise focused around qualitative methods, tool development, psychological well-being in children, health behaviour change in children, PE, PA and motor learning and development took part in a series of interactive meetings to co-produce the motivation tool. All members of the research team had at least 15 years of experience working or researching with children (maximum of 30 years). All but one had published within the SDT area, with half having published at least four SDT-related journal articles. Guidelines from COSMIN also state that the target population should be involved with the development of tools that measure an outcome within its population (Wiering et al., 2017). Therefore, development, testing and redesigning of the MAT-PE involved members of the research team working with a convenience sample of 43 children aged 5-6 years old (54% male) from three primary schools over three weeks (Supplementary Material B). This process resulted in the final version of the MAT-PE that was

deemed by the research team, through their respective relevant expertise, to be feasible and show promise of content validity that warranted further study. The MAT-PE tool and content validity testing are described in the following sections.

Description of the MAT-PE

The MAT-PE was developed as a pragmatic, mixed-method tool to overcome the challenges of conducting research with young children (Evans & Fuller, 1996, 1998) and to enable richer insights to be captured surrounding children's interpretations of their experiences (Caruth, 2013; Ponce & Pagán-Maldonado, 2015). The tool aims to assess *what* (quantitative) children's motivational perceptions are within PE and *why* (qualitative) they have those particular perceptions. The reasons for mixing the quantitative and qualitative strands within the tool was to *answer different research questions* (what and why), provide an *explanation* (qualitative to explain quantitative findings) and *illustrate* (qualitative putting 'meat on the bones' of quantitative findings) children's motivations within PE (Bryman, 2006). These aspects are depicted in Supplementary Material C.

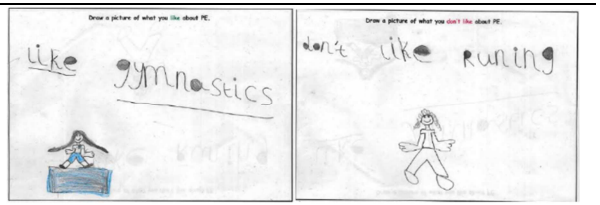
Table 2 describes the MAT-PE tool. The MAT-PE comprises a classroom draw and write activity followed by a semi-structured interview that is administered in a one-to-one format by a trained researcher. The semi-structured interview utilises a pictorial instrument and consists of interactive activities (e.g., choosing, sorting) designed to capture motivational perceptions within SDT-related constructs: enjoyment, relatedness, autonomy, competence and self-determined motivation. The use of visual resources was designed to overcome issues with children's attention span, verbal ability and abstract thinking. For each activity, the child is presented with the associated picture cards and receives a scripted set of explanations and questions from the interviewer that are compiled in the instruction manual. Children are directed to choose the card(s) that best represents their thinking (fixed choice: quantitative strand, the what) and then the interviewer asks a series of open-ended questions with probing to understand their fixed choice selection (qualitative strand, the why).

DEVELOPMENT AND VALIDITY OF THE MAT-PE

Enjoyment can be considered as an aspect of intrinsic motivation (Deci & Ryan, 1991); however, enjoyment can be seen as a standalone construct (Kimiecik & Harris, 1996), and has related positively to actual PA, PA intention, and high levels of motivation (Best et al., 2017; Bungum et al., 2000; Yli-Piipari et al., 2009). The draw and write technique was used to assess children's enjoyment of PE and was conducted as a classroom-based activity. This activity was informed by the Write, Draw, Show and Tell procedure by Porcellato et al. (2005) and Noonan et al. (2016). Children were asked to draw a picture of 'what they like about PE' on one side of an A4 blank paper and 'what they don't like about PE' on the other side.

Table 2

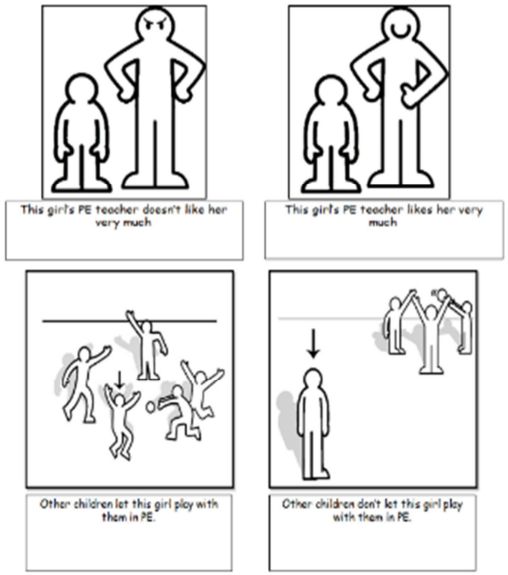
Description of the MAT-PE

Construct	Activity description		MAT-PE resources
	Whole-class activity		
Enjoyment part 1: Draw and Write	Children were given 30 minutes to draw pictures of what they liked and/or disliked about PE.		
	Activities completed one to one with researcher		
Icebreaker: Pair-matching card game	A set of PE-themed cards were laid face-up before the child. The child is asked to remember where all the matching pictures are so when turned over, they turn over only the matching pictures.		
Enjoyment part 2: Discussion around like/dislike of PE drawings	Children presented with their drawings about what they liked and/or disliked about PE.		Draw and write pictures from Part 1
	Quantitative: *I asked you to draw a picture of what you like about PE, what have you drawn here?*	Qualitative: Why do you like…?	
		Why don't you like…?	
	I asked you to draw a picture of what you don't like about PE, what you have drawn here? This is considered quantitative as children either drew/wrote what they liked or disliked, or they did not.	*You haven't drawn anything, why is that?*	

DEVELOPMENT AND VALIDITY OF THE MAT-PE

Relatedness: Choose and discuss

Children presented with two sets of two cards: one set focused on the PE teacher relationship and one set on peer relationships.



Quantitative: *This girl/boy's PE teacher likes them very much, this girl/boy's PE teacher doesn't like them very much, which girl/boy are you most like?*

Qualitative: *How do you know your PE teacher likes/doesn't like you? What do they say or do that makes you think that they like/don't like you?*

Do you like your PE teacher?

Why do you/don't you like your PE teacher?

Other children let this girl/boy play with them in PE, other children don't let this girl/boy play with them in PE, which girl/boy are you most like?

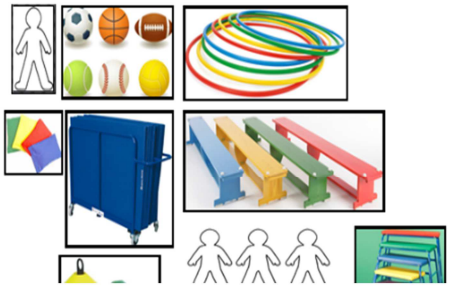
Can you tell me about a time when other children let you/didn't let you play with them in PE?

Do you let other children play with you in PE?

Is it important to let them play? Why? Why not?

Autonomy: Sorting

The child was presented with two plates: labelled "You" (the child's plate) and labelled "PE teacher" (the PE teacher's plate). Each child is shown a series of PE equipment they might be able to choose in PE and asked to sort them into whether they think they get to choose or the PE teacher chooses for them.



Quantitative: *There are some things in PE that you might get to choose and there are some things in PE that your PE might choose for you, which things do you get to choose?*

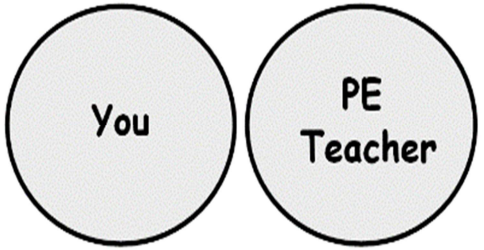
Do you ever get to choose the activities in PE or does the PE teacher?

Do you get to choose how you do movements and actions in PE or does the PE teacher show you and tell you how to do them?

If you have a question for your PE teacher, do they answer it?

If you have something to say to your PE teacher, do they listen to you?

Qualitative: *Can you tell me about a time you got to choose that?*

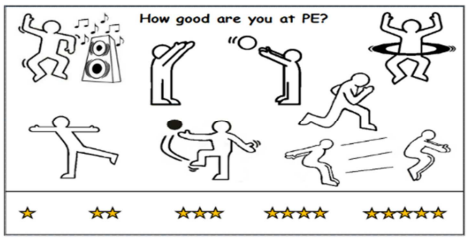


Competence: Choose and discuss

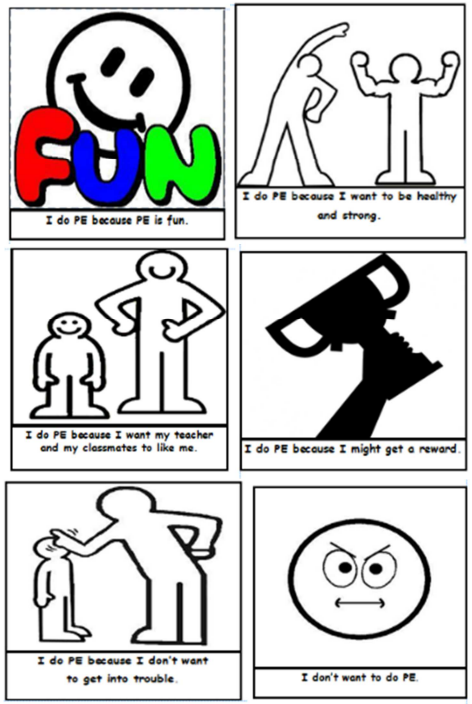
The child was presented with a series of fundamental movement skills and 1 to 5-star star-chart and were told: A child who can do all of these things all of the time would get five stars. A child who can do most of these things most of the time would get four stars. A child who can do some of these things some of the time would get three stars. A child who can a couple of things would get two stars. A child who can maybe one thing would get one star.

Quantitative: *How many stars would you give yourself for doing things in PE?*

Qualitative: *Why would you give yourself…star(s)?*



DEVELOPMENT AND VALIDITY OF THE MAT-PE

Self-regulation: Choose, sort and discuss	The child was presented with all the reasons why they might take part in PE: I do PE because PE is fun (intrinsic), I do PE because I want to be healthy and strong (identified), I do PE because I want my teacher and classmates to like me (introjected), I do PE because I might get a reward (external approach), I do PE because I don't want to get into trouble (external avoid), I don't want to do PE (amotivation). They were asked to choose their favourite reasons for taking part. They were then asked follow-up questions for each chosen reason They were then asked to place the chosen reasons in order of importance for them.		
	Quantitative: *Out of all these reasons, which are your favourite reasons for doing PE?* *Can you place your reasons into order of importance where the first means the most important?*	Qualitative: Intrinsic: *Why is PE fun?* Identified*: Why is being healthy and strong important to you?* Introjected: *Why is it important that your teacher and classmates like you? Do you ever feel like you need to do PE to show other children and teacher how good you are at PE?* External (reward): *Do you get rewards in PE? What rewards do you get in PE?* External (punishment): *If you knew you wouldn't get into trouble, would you still want to do PE? Why?* Amotivation: *Why don't you want to do PE?*	

Following completion of the draw and write activity, a trained researcher escorted the child to an adjacent location away from the classroom for the completion of the one-to-one interview. The interview commenced with a PE-themed pair-matching card game to build rapport between the child and researcher (Irwin & Johnson, 2005). Each child was then presented with their drawing from the classroom-based activity and a discussion occurred between the researcher and child about their pictures (Noonan et al., 2016; Porcellato et al., 2005).

The MAT-PE pictorial instrument and interactive activities were subsequently utilised in the interview with the child to assess each SDT construct (relatedness, autonomy, competence need satisfaction and behavioral regulation). For relatedness, questions addressed both PE teachers and peers as it has been found that both social agents effect children's relatedness (Vasconellos et al. 2019). A structured alternative response format (Barnett et al. 2015; Harter & Pike, 1984) was used (see Table 2); once the child chose which child they are most like their choice was discussed with them.

The autonomy activity focused upon the choicefulness element of autonomy, more specifically procedural (e.g., choice of equipment), organisational (e.g., peer selection) and cognitive (e.g., choice of activities; Stefanou et al. 2004). For example, children were shown a selection of PE equipment and two plates labelled "you" for the child and "PE teacher" for their PE teacher. The children were asked to sort the PE equipment onto their plate if they ever got to choose it in PE or sort onto their PE teacher's plate if the PE teacher chose it. Children were then asked to expand. Additionally, children were asked if they got to choose the movements or activities that they did in PE and if they felt that their PE teachers listened to them and answered their questions. This item centred on opportunities for input, which is considered as an autonomy characteristic (Ryan & Deci, 2017).

Within the competence activity, children were asked to rate themselves on a 1-5-star star-chart ("How good are you at things in PE?") based on pictures of fundamental movement skills which the development of is a main outcome for PE (Department of Education, 2013; UNESCO, 2013).

Within the behavioral regulation activity, each child was presented with six picture cards each representing a source of behavioral regulation (see Table 2). The pictures included a written stem that was informed by previous literature (Guay et al. 2010; Sebire et al. 2013). Integration was omitted as it is thought that this type of regulation does not emerge until adolescence or adulthood (Ryan & Deci, 2017). External regulation was split into approach (reward) and avoidance (punishment). Each type of regulation was presented to the children, one at a time, and read aloud. Children were asked to choose their most favourite reasons for taking part. For any choice they make they were then asked a related follow-up question for that type of regulation. They were then asked to put the chosen regulations into order of importance from most important to least important where more than one type of regulation can be positioned the same, e.g., intrinsic and external reward as first, external punishment as second and identified as third. Once completed, the researcher thanked the participant, gave them a sticker and escorted them back to the classroom.

Interviews were recorded using a Dictaphone; children wore microphone clips to aid recording quality. Conversations were typed up verbatim in the form of an interview transcript (qualitative) and fixed choice item selections were recorded (quantitative) for subsequent analysis (see Study 2). The total time for administration was approximately one hour, inclusive 30 minutes for the write and draw enjoyment activity and approximately 25 minutes for the SDT MAT-PE activities.

Content validity of the MAT-PE

Content validity testing of the MAT-PE was undertaken by the research team in a sample of children during baseline assessments of the cluster RCT examining PE in primary school-aged children (Rudd et al., 2020). Following recommendations from Dunn et al. (1999), content validity was also examined among researchers with expertise in SDT who were independent of the tool development.

Methods

Participants

Children

Informed written head teacher and parent/guardian consent and child assent were obtained for n=360 children from 18 Year 1 classes (5-6 years) within 12 primary schools located within a large city in North West England to participate in the cluster-RCT. A random sub-sample of eighty-five children (aged 5-6, 47% male) - approximately 5 children per class - were selected from a pool of research participants to undertake MAT-PE. These children were deemed by the class teacher to be able to speak and listen in English to an adult visitor to the school (i.e., visiting researcher).

Independent researchers

Fifteen researchers who worked within the area of SDT were contacted via email through snowball sampling; nine researchers agreed to participate in the study. This sample constituted of professors, assistant professors and lecturers in health psychology, sport and exercise psychology, and sport and movement education. Primary areas of expertise included health psychology, motor development, motivation and behaviour, exercise motivation, PE, STD, and behaviour change. This sample included a range of experience working with children (0-17 years), and within SDT (4-21 years). All but one had published within the SDT area with a range from one to 32 SDT-related publications.

Procedure

The content of the MAT-PE tool is outlined in Table 2. Following training by the lead author, a postgraduate student as well as the lead author administered the MAT-PE. Training lasted one hour and covered all aspects of tool administration including the administration script, the assessment process, activities and resources. The postgraduate student completed administration with two children under the observation of the first author before administering the MAT-PE independently. Throughout trialling the MAT-PE with the 85 children, the research team met every week over the 6-week data collection period to discuss the tool's content validity. Discussions were noted by the lead author and guided by COSMIN considerations around content validity (Terwee et al. 2018) and reviewed the relevancy (were the questions relevant to the construct?),

comprehensiveness (was each aspect supported conceptually in accordance with the theoretical framework?) and comprehensibility of the activities (did the children understand the activities and what they were supposed to do?).

Independent SDT researchers were sent a matching task (Hambleton, 1980 in Dunn et al., 1999) via email to ascertain content validity. In order to complete the task, researchers had to match each question/stem with the corresponding construct (enjoyment, relatedness, autonomy, competence, intrinsic, identified, introjection, external approach, external avoid and amotivation). They were also asked, on a scale from one to five, to rate each item on how relevant (1=poor match, 2=fair, 3=good, 4=very good, 5=excellent match) and comprehensive (1=poor comprehensiveness, 2=fair, 3=good, 4=very good, 5=excellent comprehensiveness) each item was within that construct. Matching scores for each item was determined through the number of researchers out of the sample who correctly aligned it with the designated construct within the MAT-PE, culminating in a percentage score. Mean scores were calculated for relevance and comprehensiveness.

Results

The MAT-PE required approximately 25 minutes to administer (not including the 30-minute classroom drawing task). The research team reached consensus that the MAT-PE elicited sufficient depth from the children according to their enjoyment of PE, their BPNS and behavioral regulation. It was decided that the tool was *relevant* as all activities were judged to include aspects pertinent to each theoretical construct and were representative of an early primary school PE context; *comprehensive* as all activities encompassed sufficient components to ensure key considerations of BPNS and behavioral regulation were assessed to the fullest extent possible in this age group (e.g. inclusion of PE teacher and peer groups within the relatedness activity; addition of follow-up questions related to cognitive autonomy, i.e. choice over activities and movement), and *comprehensible* as activities were familiar and understood by the children (e.g. drawing, sorting, choosing). Therefore, consensus among the research team, via COSMIN guidelines, that content validity was reached.

Table 3 shows the content validity matching task results. Eleven of the 19 questions/stems were matched to the corresponding construct by at least 75% of the researchers (items 3, 4, 5, 7, 8, 12, 13, 14, 17, 18, and 19); five of which were matched by 100% of the researchers (items 7, 8, 17, 18, and 19); four items by around half of the researchers (55.55-66.66%; items 2, 6, 9, and 16), and four items by a third of the researchers or less (11.11-33.33%; items 1, 10, 11, and 15). The majority of items were judged to be "good" (score of 3) or above for relevance and comprehensiveness. Item 1 was matched by a third of the researchers and item 11 by one researcher, however, both items were judged as "very good" on relevance and comprehensiveness. Item 10 was also matched by a third of the researchers, however, it was judged as "good" on relevance and "fair" on comprehensiveness. Stem 15 was matched by two researchers and judged to have "excellent" relevance and "good" comprehensiveness.

Table 3

Matching percentage, relevance and comprehensiveness of the MAT-PE items and their constructs

Construct		Question/Stem	Matching (%)	Relevance Mean (SD)	Comprehensiveness Mean (SD)
Enjoyment	1.	Like PE	33.33	4.67 (.58)	4.50 (1.00)
	2.	Dislike PE	55.55	4.25 (.50)	4.25 (.96)
Relatedness	3.	Liked/Disliked by PE teacher	88.88	4.00 (1.07)	3.43 (1.13)
	4.	Like/Dislike of PE teacher	88.88	3.62 (1.06)	4.17 (.98)
	5.	Included/Excluded by peers	77.77	3.29 (1.11)	3.00 (1.09)
	6.	Includes/Excludes peers	66.66	2.29 (.95)	2.71 (1.38)
Autonomy	7.	PE equipment choice	100	4.56 (.73)	3.88 (1.55)
	8.	Choice of movements	100	4.22 (1.09)	3.88 (1.55)
	9.	Choice of activities	66.66	4.50 (.84)	4.00 (1.73)
	10.	Listened to by PE teacher	33.33	3.67 (.58)	2.67 (1.53)
	11.	PE teacher answers questions	11.11	4.00*	4.00*
Competence	12.	Self-rating of FMS	88.88	4.11 (1.27)	3.86 (1.07)
Intrinsic	13.	I do PE because it's fun	88.88	4.63 (1.06)	4.83 (.41)
Identified	14.	I do PE because I want to be healthy and strong	77.77	4.29 (1.25)	4.00 (1.55)
Introjected	15.	I do PE because I want my PE teacher and classmates to like me	22.22	5.00 (.00)	3.50 (2.12)
Introjected	16.	Do you ever feel like you need to do PE to show other children and your teacher how good you are PE?	55.55	4.40 (.89)	4.25 (1.50)
External approach	17.	I do PE because I might get a reward	100	4.44 (.73)	4.63 (.74)
External avoid	18.	I do PE because I don't want to get into trouble	100	4.88 (.35)	3.86 (1.68)
Amotivation	19.	I don't want to do PE	100	4.67 (.71)	4.25 (1.16)

Note. SD = Standard Deviation, * = data from one person therefore Standard Deviation could not be computed for that item

Study 2: Development, content validity, acceptability and reliability of the MAT-PE codebook

Study 2 was concerned with developing an approach to enable the mixing of quantitative (fixed choice selection) and qualitative (open-ended question responses) MAT-PE data for analysis (Creswell & Plano Clark, 2011). In the present study, the quantitative strand took priority as the qualitative strand helped to explain and illustrate the quantitative data (Bryman, 2006). Furthermore, a quantitative priority for analysis was sought in order to facilitate the statistical investigation of motivational profiles, the antecedents and consequences of motivation, and to provide numerical data that could be analysed in longitudinal and experimental research. Thus, quantitative content analysis (Rourke & Anderson, 2004) was selected as this is an acceptable form of deductive analysis for semi-structured interviews and can be used to count the frequency and intensity of responses. An important stage of quantitative content analysis is to establish a coding scheme that allows testing of hypotheses (Rourke & Anderson, 2004; White & Marsh, 2006). Therefore, Study 2 aimed to develop a 'codebook' for researchers so that the transcript data from the MAT-PE could be analysed by coding young children's motivational perceptions towards PE, mixing the quantitative and qualitative strands. Furthermore, this study aimed to examine the content validity and acceptability of the developed codebook, and determine inter-rater and intra-rater reliability.

Development of the MAT-PE codebook

Six members of the research team (KFD, PW, JR, SR, FB, LF) from the MAT-PE development were involved in creating the codebook and provided the necessary skill, labour, thinking and energy (Fernald & Duclos, 2005). Following previous research (Fonteyn et al., 2008; MacQueen et al., 1998), the codebook was developed through an iterative process and structured similarly. The research team met on six occasions over a three-month period to review and refine the codebook. This included confirming codes, determining coding, and checking for ambiguous wording in code descriptions. In the final step, four of the research team (KFD, JR, SR, LF) coded the same transcript

data and found few discrepancies in coding. Thus, consensus was reached among the research team that the codebook development process was complete.

The final MAT-PE codebook (Supplementary Material D) was scaffolded and underpinned by SDT and included codes (numerical), code descriptions and code examples. A coding table was included with predetermined categories for each construct within the MAT-PE: enjoyment, relatedness, autonomy, competence need satisfactions and behavioral regulation. Codes for each motivational construct were initially created by reading through randomly selected transcript data from Study 1.

Codes were numerical, whereby higher values indicated higher levels of motivational perceptions. This quantitative content analysis (White & Marsh, 2006) approach was used in order to understand and describe motivational perceptions in a way that can be counted, quantified and therefore measured. The numerical scoring process was designed to take into account the child's initial quantitative response/choice (the 'what': yes or no, this or that) alongside the qualitative nature of the child's answer (the 'why'), and whether the child provided a surface level (gave no more detail to their initial answer) or deep level response (gave more detail to their initial answer) to the researcher's questions. Deep level responses were taken to indicate stronger motivational perceptions whereas surface level responses were taken to indicate weaker motivational perceptions. Positive and negative aspects of each construct were therefore merged within the same coding matrix. For example, in the relatedness activity children chose between being included or excluded by peers in PE. Responses were put on the same coding scale from the most negative (scored 1: excluded, deep level response) to most positive (scored 4: included, deep level response). Code descriptions outlined the choice and depth of response for each code, while code examples included direct quotes from children's actual transcript data, providing authenticity. Examples of coding for a child who picked a positive option and gave a deep-level response for the relatedness activity can be seen in Table 4.

Overall construct scoring differed by construct: enjoyment score was calculated by taking the coding given in “Like of PE” and subtracting the coding given in “Dislike of PE” which provided a range from -3 to +3. Codes from all four relatedness responses were added together to create the overall relatedness score, giving a range from 4 to 16. The same was done for autonomy where all four responses were added to create the overall autonomy score, giving a range from 4 to 15. Competence included one item only and therefore constituted the overall score (1-9). For autonomous motivation, the coding given for intrinsic and identified regulations were added and then divided by two to obtain the mean. For controlled motivation, first the coding for external regulations (approach and avoidance) were added and then divided by two to obtain a mean.

Table 4

An example of coding from the MAT-PE codebook

Relatedness Satisfaction| Liked/Disliked by PE Teacher | Activity 2a

Question(s): *This girl's/boy's PE teacher likes her very much, this girl's/boy's PE teacher doesn't like her very much, which girl/boy are you most like?*
Follow-up question(s): *How do you know your teacher likes/doesn't like you?*

Code	Description	Example
4✓	The child has chosen "liked by teacher" and articulates a deep level response as to how they know that.	R: "How do you know your PE teacher likes you?" → C: "Because sometimes he says good work" → C: "Because she never gets angry at me and she lets me help her" → C: "Because I do good work."
3	The child has chosen "liked by teacher" and articulates a surface level or irrelevant response as to how they know that.	R: "How do you know your PE teacher likes you?" → C: "They just do." → C: "Everyone is supposed to like everyone." → C: "Because I like ice cream." → C: "I don't know."
2	The child has chosen "disliked by teacher" and articulates a surface level or irrelevant response as to how they know that.	R: "How do you know your PE teacher doesn't like you?" → C: "I don't know." → C: "Because I like ice cream." → C: "I don't know."
1	The child has chosen "disliked by teacher" and articulates a deep level response as to how they know that.	R: "How do you know your PE teacher doesn't like you?" → C: "Because he be mean to me" → C: "Because sometimes he says I'm naughty."
N/A	The child has **failed to choose** between the two options and has not articulated toward which choice they feel more affinity with when prompted by the researcher.	The child may choose both or neither to obtain an N/A.

Coder's comments (e.g. if they provided an irrelevant response, any notable comments):
"A: Erm, because she smiles at me all the time… She doesn't pull angry faces at me"

Child has described **how** they know the PE teacher likes them and therefore the response is considered a deep-level response.

This mean was then added to introjection and then divided by two to obtain a mean for overall controlled motivation. Amotivation included one item only and therefore constituted the overall score.

Content validity and acceptability of the MAT-PE codebook

Methods

Participants

Four individuals (50% female) with a range of SDT experience, who were independent of the research team, were asked to use the codebook to code a transcript from Study 1. Two of the coders were academics in psychology/sport coaching with 10-11 years of experience in their area of interest. The other two coders were post-graduate psychology students with 4 to 6 years in their area of study.

Procedure

A brief explanation of how the MAT-PE is administered was given to participants before being asked to read the instruction manual. Participants were given time to code the designated transcript and were asked to note down any thoughts or queries that they had whilst using the codebook so that they would not have to rely on recall. Each participant was asked a series of content validity and acceptability questions regarding each part of the codebook. Content validity questions referred to: relevance (*Is the code table relevant for the construct of interest? Are all code options independent of each other with no overlapping or ambiguous descriptions and examples?)* comprehensivenesss *(Are there any key concepts not covered by the codes?)* and comprehensibility *(Are the instructions understandable? Is the language used in the code table understandable?)* (Terwee et al, 2018). Acceptability questions (*Were any sections difficult to complete? Would you change anything in the code table to improve it?)* investigated the codebook's appropriateness. Responses were captured through participants writing their answers to each question followed by a discussion between the researcher and participants, which was recorded via Dictaphone. Written

answers were inputted into a spreadsheet and the lead author listened to the recorded discussions and added any extra information, which was provided verbally, into the spreadsheet.

Results

Coding took approximately 30 minutes (including reading of transcript, allocation of codes). All four individuals who completed the codebook content validity and acceptability agreed that for each construct (enjoyment, relatedness, autonomy, competence and self-regulation) the codebook was relevant, comprehensive and comprehensible. When asked if they had any recommended changes that would make the codebook easier to use, the feedback included: to provide more examples (enjoyment), put in place a way to keep track of the chosen equipment (autonomy), and to label the type of motivation in the instruction booklet (behavioral regulation). These recommendations were taken on board and the codebook was amended.

Inter-rater and intra-rater reliability of the MAT-PE codebook

Methods

Participants

Three individuals (100% female) with SDT knowledge were asked to determine inter-rater reliability of the codebook. These individuals consisted of a post-graduate student who had helped determine acceptability of the codebook, an academic and researcher in the area of psychology and SDT (second author) and the first author.

Measures and procedure

To determine inter-rater reliability, each individual was given the codebook, the instruction manual and eight transcripts from eight different children provided through the MAT-PE tool. Transcript data consisted of verbatim responses from children collected during the MAT-PE administration. Transcripts were randomly selected via a computerised number generator to include four from Study 1 and four from a later time point (Study 3). Intra-rater reliability was

examined by investigating the consistency between codes when the same eight transcripts were coded by the first author on two separate occasions one week apart.

Data analysis

Statistical tests were completed using SPSS, version 24 [IBM SPSS Statistics Inc., Chicago, IL, USA]. For inter-rater and intra-rater reliability (IRR), intraclass correlation coefficients (ICC), two-way mixed single measures for absolute agreement with 95% confidence intervals (95% CI), were used to determine the level of agreement between three raters (inter-rater reliability) and between two time points (intra-rater reliability). The IRR was interpreted with cut-offs set at less than 0.40 (poor), 0.40 to 0.59 (fair), 0.60 to 0.74 (good) and 0.75 to 1.0 (excellent; Cicchetti, 1994).

Results

Inter-rater reliability for PE enjoyment, relatedness, autonomy, competence, autonomous motivation, controlled motivation all had an ICC above 0.9, which is considered excellent (Cicchetti, 1994). As there was zero variance in the coding for amotivation for all eight transcripts, no ICC could be calculated for this construct. However, the scores had 100% agreement between the three raters.

Intra-rater reliability for PE enjoyment, relatedness, autonomy, competence, autonomous motivation, controlled motivation all had an ICC above 0.9, which is considered excellent (Cicchetti, 1994). As there was zero variance in the coding for amotivation for all eight transcripts, SPSS could not generate an ICC for this construct, however the scores had 100% agreement from test-to-retest.

Study 3: Construct validity of the MAT-PE

Construct validity is an important measurement property for evaluating outcome measures (Prinsen et al., 2016). Study 3 therefore aimed to explore the construct validity of the MAT-PE through hypothesis testing using correlational data. Following SDT research (Ryan & Deci, 2017; Sebire et al., 2013; van Aert et al., 2017), broadly we hypothesised that 1) BPN will positively associate with each other, 2) BPN will positively associate with autonomous regulations and

negatively associate with controlled regulations and amotivation, 3) introjection will demonstrate complex associations with the other variables, and 4) behavioral regulations will ascribe to the simplex model (Ryan & Connell, 1989). Descriptive data is also presented to illustrate the data that can be collected from the MAT-PE and its codebook. Data was collected during post-test assessments of the cluster-RCT (Rudd et al., 2020).

Method

Participants

Participants involved in Study 1 also formed a convenience sample for this study. Seventy-eight children (male=48.7%, White British=57.7%, age_m=6.34 years, SD=0.30) took part. Seven children from Study 1 did not take part in this study due to being absent on assessment days or leaving school.

Measures and Procedure

MAT-PE was used with each child in accordance with the procedures outlined in Study 1. The codebook developed in Study 2 was used to code the transcript data obtained from the 78 children. The first author and two trained postgraduate students administered the MAT-PE with children and the first author coded the data with the codebook (ICC above 0.9).

Data analysis

All statistical tests were completed using SPSS, version 24 [IBM SPSS Statistics Inc., Chicago, IL, USA]. The numerical codes for each theoretical construct, assigned using quantitative content analysis as outlined in the MAT-PE codebook, were used in data analysis (higher numerical codes represented stronger motivational perceptions). Descriptive statistics were computed for the overall sample. Median values and inter-quartile ranges were used for descriptives due to the categorical nature of the data. Wilcoxon signed rank tests were conducted to investigate differences in behavioral regulation choices. A Kendall's tau-b correlation was run to determine the relationships between BPNS, behavioral regulation and enjoyment as captured by the MAT-PE (Table 7). Kendall's statistic was used due to the small sample size and it is said to be a better estimate of the correlation

in the population in comparison to Spearman's statistic (Howell, 1997). Following COSMIN guidance, the results describe the direction and magnitude of relationships and avoided reporting p values, which are affected by sample size (Mokkink et al., 2012).

Results

The MAT-PE descriptive data for the overall sample is presented in Table 5. The median, inter-quartile ranges and minimum and maximum scores indicate that children chose a variety of responses, demonstrating that all choices were valid.

Table 5

Descriptive statistics for N=78 children on the MAT-PE Version 1.

Code construct (PSR)	Min	Max	Median (IQR)
Enjoyment (-3 to +3)	-1	+3	0.00 (0.00,3.00)
Like PE (1-4)	2	4	4.00 (4.00,4.00)
Dislike PE (1-4)	1	4	4.00 (1.00,4.00)
BPNS Total (9-40)	25	39	35.00 (34.00,37.00)
Relatedness (4-16)	11	16	15.00 (15.00,16.00)
Liked by PE teacher (1-4)	3	4	4.00 (4.00,4.00)
Like of teacher (1-4)	3	4	4.00 (4.00,4.00)
Inclusion by peers (1-4)	1	4	3.50 (3.00, 4.00)
Inclusion of peers (1-4)	1	4	4.00 (4.00,4.00)
Autonomy (4-15)	7	15	11.00 (11.00,13.00)
*Pictoria*l (1-6)	2	6	4.00 (4.00,6.00)
Move/activities (1-3)	1	3	1.00 (1.00,2.00)
Listened to (1-3)	1	3	3.00 (3.00,3.00)
Questions answered (1-3)	1	3	3.00 (3.00,3.00)
Competence (1-9)	2	9	9.00 (8.00,9.00)
Autonomous (1 to 5)	1	5	3.50 (3.00,4.00)
Intrinsic (1-5)	1	5	3 .00 (3.00,5.00)
Identified (1-5)	1	5	3.50 (3.00,5.00)
Controlled (1 to 5)	1	4.5	2.25 (1.50,2.75)
External reward (1-5)	1	5	3.00 (2.00,4.25)
External punishment (1-5)	1	4	1.00 (1.00,2.00)
Introjection (1-5)	1	5	2 .00 (1.00,3.00)
Amotivation (1-5)	1	5	1.00 (1.00,1.00)

Note. PSR = Possible Score Range, BPNS = Basic Psychological Needs Satisfaction, Min = Minimum, Max = Maximum, IQR = Inter Quartile Range. Autonomous and controlled motivation scores are mean scores of the sub-constructs within them (e.g., Autonomous motivation = (intrinsic + identified)/2) with higher scores indicating stronger motivation

Enjoyment

For enjoyment, while the group median value was 0, the interquartile range (IQR) indicates that 75% of coding fell between 0 and 3 (maximum score), signifying that overall, the majority of children enjoy PE to a greater extent than they dislike PE.

Basic Psychological Need Satisfaction

For the overall sample, the median value was 35 with 75% of coding between 34 and over (maximum score 39). Higher coding in the majority of the sample of this summary construct indicate that overall, all three basic psychological needs are highly satisfied. The overall competence and relatedness median codes and IQR indicated that these basic psychological needs were highly satisfied within the majority of the sample. The overall autonomy median and IQR values indicate that the basic psychological need for autonomy was moderately satisfied within 50% of the sample, and highly satisfied in 25% of the sample (see Table 5). High levels of procedural and organisational autonomy need satisfaction were found in the majority of the sample for *choice of equipment and peers* (pair- and/or group-work) in the pictorial activity. Lower median codes and IQR values were found for cognitive autonomy need satisfaction in terms of choice of movement/activities in PE with 75% of coding falling at 2 and under (maximum of 3).

Behavioral regulation

As shown in Table 6, the most popular behavioral regulations for taking part in PE were intrinsic, identified and external reward (87.17%, 84.62%, 79.49% respectively) with introjected and external punishment as less popular behavioral regulations for taking part in PE (66.67% and 33.33% respectively). The least chosen was amotivation (2.56%). At *least* a third of the sample ranked an autonomous form of motivation as their first choice and at *most* a third of the sample chose a controlled form of motivation as their first choice for taking part in PE.

Table 6

Number and percentage of behavioral regulations chosen overall, as first choice, as other choice and not

picked by children

Type of regulation	No. of children (Total$_n$=78)	1st choice	"other" choice	Not Picked
Autonomous Motivation				
Intrinsic	68 (87.18%)	26 (33.33%)	42 (53.85%)	10 (12.82%)
Identified	66 (84.62%)	39 (50.00%)	27 (34.62%)	12 (15.38%)
Controlled Motivation				
Introjected	52 (66.67%)	6 (7.69%)	46 (58.97%)	26 (33.33%)
External reward	62 (79.49%)	23 (29.49%)	39 (50.00%)	16 (20.51%)
External punishment	26 (33.33%)	3 (3.85%)	23 (29.49%)	52 (66.66%)
Amotivation	2 (2.56%)	1 (1.28%)	1 (1.28%)	76 (97.44%)

High percentages of children gave deep level responses (verbally provided more detail to their fixed choice selection) for choosing amotivation, external reward, intrinsic and identified (100%, 85.48%, 85.29% and 84.85% respectively) reasons. A reasonable number of children provided deep level responses for introjection (65.38%) while less than half gave deep level responses to external punishment (48%). Children had six types of behavioral regulations to choose from and their number of choices varied: 35.90% chose four regulation types, 23.80% chose five and 20.51% chose three regulation types, 12.82% chose two types of regulations, and 6.41% chose one type of regulation. No child chose all regulation types, and this variance in choices demonstrates that children can differentiate between the different types of regulations, as well as being able to provide deep level responses. To view a Figure that shows the variance in the number of behavioral regulations across the sample, please see Supplementary Material E.

Autonomous and Controlled motivation

The overall autonomous motivation median and IQR values indicate that the majority of children were experiencing moderately high levels of autonomous motivation. The overall controlled motivation median and IQR values indicate that the majority of children were experiencing low to moderate levels of controlled motivation in PE. A Wilcoxon signed-rank test showed a statistical difference between external positive regulation and external negative regulation ($Z=-6.69$, $p<0.001$), external positive regulation and introjected regulation ($Z=-3.94$, p,0.001), and external negative and introjection ($Z=-5.21$, $p<0.001$). This indicates that the types of controlled regulations were chosen and responded to differentially.

Amotivation

The overall amotivation median and IQR values indicate that although amotivation is very low in this sample, it is still present.

Construct validity: Hypothesis testing

Correlational data for the MAT-PE is presented in Table 7.

Table 7

Kendall's tau-b Correlational Matrix between Enjoyment, Basic Psychological Needs, and Behavioral Regulation Constructs captured by the MAT-PE.

N = 60	1	2	3	4	5	6	7	8	9	10
1. Enjoyment	-									
2. Relatedness	.07 (-.16 - .30)	-								
3. Autonomy	.04 (-.20 - .29)	.24 (.04 - .41)	-							
4. Competence	-.04 (-.27 - .21)	.23 (-.01 - .45)	.14 (-.11 - .36)	-						
5. Intrinsic	.12 (-.10 - .35)	.13 (-.11 - .38)	.02 (-.23 - .27)	.09 (-.17 - .35)	-					
6. Identified	-.09 (-.31 - .14)	.01 (-.23 - .25)	.01 (-.22 - .24)	.09 (-.16 - .32)	-.24 (-.50 - .00)	-				
7. External app.	.12 (-.11 - .36)	.11 (-.12 - .32)	.04 (-.22 - .29)	.29 (.08 - .49)	-.05 (-.26 - .18)	-.08 (-.32 - .15)	-			
8. External avo.	-.04 (-.30 - .20)	.16 (-.09 - .39)	.05 (-.19 - .24)	.00 (-.23 - .22)	.17 (-.10 - .42)	.12 (-.10 - .33)	.24 (.04 - .44)	-		
9. Introjected	.08 (-.13 - .31)	.09 (-.15 -.31)	.24 (.03 - .42)	.12 (-.13 - .35)	.19 (-.04 - .44)	.15 (-.16 - .41)	.28 (.09 - .48)	.40 (.18 - .60)	-	
10. Amotivation	-.12 (-.24 - -.05)	-.14 (-.30 - -.01)	-.16 (-.30 - -.06)	-.21 (-.38 - -.09)	-.27 (-.43 - -.17)	-.04 (-.32 - .22)	-.13 (-.29 - .02)	-.12 (-.22 - -.07)	-.19 (-.32 - -.12)	-

Note. Brackets include bias corrected accelerated confidence intervals set at 95% with 1000 bootstraps; app. = approach, avo. = avoidance.

Hypothesised/Expected relationships

Small but positive associations were found between all three BPN (.14 to .24). There were also a small, positive association between relatedness need satisfaction and intrinsic regulation (.13). All other associations between BPNS and the autonomous types of regulations (intrinsic and identified) were weak and under .10, although all were in the expected direction (positive). Introjection had a small, positive association with identified regulation (.15). Introjection also had a small, positive association with external (approach; .28); however, introjection had a stronger, positive association with external (avoidance; .40).

There was a small, negative association between intrinsic regulation and amotivation (-.27). Amotivation had weak to small, negative associations with enjoyment, all three BPN, intrinsic and identified regulation (-.04 to -.27). The autonomous types of regulation and external (approach) had negative associations under .10; however, associations were in the expected direction.

Unexpected relationships

Relatedness needs satisfaction had small, positive associations with both external approach and avoidance (.11 and .16, respectively). There was a small, positive association between competence and external (approach; .29); however, there was no relationship between competence and external (avoidance; .00). Autonomy need satisfaction and external (approach and avoidance) regulation had associations below .10; however, the associations were in an unexpected direction (positive).

A small, negative association between intrinsic and identified regulations (-.24) was found. Intrinsic regulation also had small, positive associations with external (avoidance; .17) and introjected regulation (.19). Identified regulation had a small, positive association with external (avoidance; .12). amotivation had small negative associations with controlled types of motivation (-.12 to -.19).

General Discussion

Despite the ability of young children to report on their own experiences (Stone & Lemanek, 1990), and their suspected ability to differentiate between motivational constructs (Butler, 2005; Guay et al., 2010), there is a distinct lack of appropriate tools to measure young children's motivation (Sebire et al., 2013), particularly within PE. This paper reported the development and content validity of the MAT-PE and its associated codebook and presented preliminary descriptives and an exploration of construct validity, as guided by COSMIN (Mokkink et al., 2010; Mokkink et al., 2012; Terwee et al., 2018). Study 1 developed the Motivation Assessment Tool for Physical Education (MAT-PE), a mixed-method, age-appropriate tool for assessing 5-6-year-old children's motivation for PE, and found the tool was judged to have promising content validity. Study 2 developed a codebook to analyse transcript data from the MAT-PE. The codebook was found to be acceptable by researchers with differing SDT experience, judged to have content validity and demonstrated excellent inter- and intra-rater reliability. Study 3 presented illustrative MAT-PE data that showed that children had high enjoyment, relatedness and competence need satisfaction and lower autonomy need satisfaction. Children also had moderate to high autonomous motivation, low to moderate controlled motivation and low amotivation. Finally, children's MAT-PE data demonstrated a mixture of expected and unexpected relationships in accordance with hypothesis testing for construct validity. The following sections provide a detailed discussion of these findings and their implications for the MAT-PE tool.

Content validity

Content validity is arguably the most important psychometric property to determine suitability of a measurement tool, as without content validity, other types of validity cannot be conducted (Prinsen et al., 2018). According to COSMIN (Terwee et al., 2018), a tool has good content validity when its items and instructions are relevant, comprehensive and comprehensible. The involvement of the target population (Wiering et al., 2017) led to a tool which comprehensively

captures BPNS and behavioral regulations in PE - including distinct assessments of introjected and external regulations which were collapsed in previous measures (Guay et al., 2010). Comprehensibility was demonstrated by the children who were able to pick from the different types of regulation, and also provide deep level responses to the follow-up questions. Follow-up questions were put in place to ascertain children's level of understanding around these different types of behavioral-regulation and informed researchers as to why they partook in PE. Content validity was further tested through independent researchers with expertise in SDT who completed a matching task and rated each MAT-PE item on its relevance and comprehensiveness. Although 11 of the 19 items were matched to the intended construct by at least 75% of the researchers, there were four items that were matched by between 55.55% and 66.66%, and four items that were matched by less than 50% of the researchers. This suggests that the MAT-PE tool shows promise of content validity and theoretical fidelity. Yet, as discussed below in relation to each construct, eight items were potentially problematic and may require further development.

Enjoyment

PE enjoyment was included in the MAT-PE as enjoyment is positively related to actual PA, PA intention and high levels of motivation (Best et al., 2017; Bungum et al., 2000; Yli-Piipari et al., 2009). "Like of PE" and "Dislike of PE" items were matched to the enjoyment construct by only 33.33% and 55.55% of the independent researchers, respectively. It is interesting that "Dislike of PE" was matched more successfully than "Like of PE". "Like of PE" might have been perceived by the independent researchers to relate to intrinsic motivation to a greater extent than enjoyment. Enjoyment is a significant part of intrinsic motivation (Deci & Ryan, 1991). Enjoyment can also be viewed as a standalone construct, which has a history of challenge in its definition and use (Kimiecik & Harris, 1996). Perhaps these findings indicate an issue with the matching task methodology, as there would be an assumed association between enjoyment and intrinsic motivation. The advantage of retaining these enjoyment items in the MAT-PE is that enjoyment of PE can be understood in its own right. If the researcher is not interested in PE enjoyment, these items could be omitted from

administration of the tool without impacting perceptions of intrinsic motivation. Removing the enjoyment items and the associated write and draw activity would also reduce MAT-PE administration time.

Relatedness

"Liked/Disliked by PE teacher" and "Like/Dislike of PE teacher" were matched highly by the independent researchers (88.88% respectively). Matching percentage was lower, however, for the peer related items ("Included/Excluded by peers" and "Includes/excludes peers"). The item which focused on children including others was matched lower than the item focused upon being included by others. It was decided to include the former item as Ryan and Deci (2017) state that "Relatedness refers to both experiencing others as responsive and sensitive and being able to be responsive and sensitive to them…" (p.86). This implies a two-way meaningful interaction between social agents and consequently it would be pertinent to retain the item. As such, it would be advantageous to include both types of questions per social agent (i.e., included/excluded *by* peers and includes/excludes *of* peers). A recent systematic review and meta-analysis reported differential effects of teachers and peers on relatedness need satisfaction (Vasconcellos et al., 2019), indicating that peer items should be retained in the MAT-PE tool.

Autonomy

Within autonomy, it was interesting that choice over equipment and movements was matched by 100% of the independent researchers while choice over activities was only matched by 66.66%. It is not clear why the activity choice item was not matched as much as the other choice items as it was phrased in the same way. The autonomy construct was limited to choice and lacked a measure of volition. Despite this lack comprehensiveness in the construct, choice is a significant factor of autonomy as demonstrated by De Meester et al. (2020) who found that providing 12-13-year-old children with choice in activity, level, pace and with whom they wanted to work with may have positively impacted competence and relatedness need satisfaction.

"Listened to by teacher" and "PE teacher answers questions" were matched by 33.33% and 11.11%, respectively. In their comments, the independent researchers mentioned that this item was around a choice of being listened to, which has been identified as an inclusive element of autonomy items in other measures (Smith et al., 2015). "Listened to by teacher" was rated lower for relevance and comprehensiveness; however, it is also closely related to feelings of inclusivity within autonomy. Despite this connection with theory, the tool may benefit in future by omitting these items in favour of items around volition, which was not included within version 1 of the MAT-PE. For example, an item could be developed based on an item from the children's version of the Basic Psychological Needs Satisfaction and Frustration Scale "I do the things I do because I really want to do them" (BPNSFS; Van der Kaap-Deeder et al., 2015).

Although the autonomy activity was predominantly based on children's perceptions of how much choice they felt during their PE lessons, it could be contended that the language used within these items lies somewhere between need satisfaction and need support. There is arguably a fine line between perceptions of need support and need satisfaction when stemming from the participant self-reporting (rather than a measure of need support via an external agent such as through observation). The items were aimed at targeting the perceptions that the children felt rather than what was being made available to them; however, it is unclear whether children of this age can tell the difference between these two types of questions. Regardless, effort should be made in future to align the wording of these MAT-PE items to reflect an inner process rather than an availability of support.

Competence

Although the competence item was highly matched by independent researchers, it would be prudent to add additional items to further investigate younger children's competence perceptions. The competence construct could be developed further by introducing other components, perhaps based on items of the BPNSFS (Van der Kaap-Deeder et al., 2015), such as, "I am good at difficult tasks", as currently competence is a global item centred on ability to perform fundamental

movement skills within version 1 of the MAT-PE, following the UK national curriculum aims for primary PE.

Behavioral regulation

All items were highly matched by independent researchers (77.77%-100%) except for items of introjection based on Guay et al.'s (2010) motivation measure. "I do PE because I want my PE teacher and classmates to like me" and "Do you ever feel like you need to do PE to show other children and your teacher how good you are at PE?" were matched by 22.22% and 55.55% of researchers, respectively. One independent researcher commented that the former item seemed more external than introjected. We would argue that the use of "I want" within the stem aims towards a more internal rather than purely external drive. Introjection has often been presented within motivation measures as a form of guilt (e.g., "I feel guilty when I don't exercise", Markland & Tobin, 2004). Sebire et al. (2013) also tapped into guilt within their measurement of motivation within PA for 7.84- to 11.09-year-old children with the stem "*When I'm not active I feel bad*". Within the current study, only one child could provide a definition approaching guilt and the use of 'feeling bad' also caused confusion. Feelings of guilt stem from relating affect to events with increasing social understanding (Malti, 2016), which only occurs through experience. As young children may not have sufficient opportunities to effectively form links between affect and events, feelings of guilt are perhaps less appropriate than feelings of ego to capture introjection. Clearly, this is a challenging construct to measure in this age group and there is a need to further develop this item. Further investigations into this age group regarding introjected regulation would be advantageous for the field and SDT.

Construct validity

Small, positive associations between BPNS were found, which is in line with previous SDT studies (Huhtiniemi et al., 2019; Ryan & Deci, 2017; Sebire et al., 2013). However, the associations in this study were smaller in comparison to these studies, but consistent with some other studies with children (e.g., van Aert et al., 2017). Although associations between BPNS and the autonomous types

of regulation (i.e., intrinsic and identified regulations) were in the expected direction (Deci & Ryan, 2000), associations were weak (<.10). Given that some of the autonomy need satisfaction items were matched lowly, and considerations of the language used (support vs. satisfaction), these content validity issues may explain some of the weak associations between autonomy and autonomous types of regulation. It is unclear why relatedness and identified regulation had a positive but weak association given that both relatedness and identified items were highly matched by independent researchers and based on past research (e.g., Guay et al., 2010; Sebire et al., 2013).

Introjection had a complicated set of associations with various constructs which aligns with the partially internalised nature of introjection (Ryan & Deci, 2017). For instance, introjection was positively associated with identified (autonomous) and external (controlled) regulations as well as BPNS (Vasconcellos et al., 2019). However, there was a difference between the different types of external regulation and introjection where external avoidance (i.e., fear of punishment) was more (positively) strongly associated in comparison to external approach (i.e., chance for a reward). This indicates that a child who seeks to avoid punishment is more likely to want to please their teacher and peers. Amotivation was negatively associated with enjoyment, BPNS and autonomous types of regulation which was to be expected (Gao et al., 2013; Vasconcellos et al., 2019).

Notably, the most unexpected association was the small, negative association between intrinsic and identified regulation. Previous research has consistently found positive associations between the two autonomous types of regulation (Gao et al., 2013; Huhtiniemi et al., 2019; Ntoumanis, 2001; Sebire et al., 2013) which ascribe to the simplex model (Ryan & Connell, 1989). The negative association found between intrinsic and identified regulations in this study indicates that the more children perceive PE to be fun (intrinsic) the less likely they are to participate to be healthy and strong (identified) and vice versa. It may be possible that children of 5- and 6-years of age perceive these types of regulation as opposing. If they perceive PE to be highly fun, they may not feel that being healthy and strong is as important and vice versa.

The separation of external regulation into approach and avoidance led to a small, positive association between external approach and competence while there was no correlation between external avoidance and competence. This is in contrast to other research which has found negative correlations between external regulation (not split) and competence (Huhtiniemi et al., 2019; Sebire et al., 2013; Vasconcellos et al., 2019), and against SDT theory which posits that controlled types of motivation are negatively associated with BPNS (Ryan & Deci, 2017). The findings in this study indicate that children who felt higher levels of competence felt that participation in PE was highly driven by external rewards and vice vera. This indication may not be so surprising as the use of rewards in education is prevalent (Deci et al., 2001) and children are highly competence driven (Harter, 1988). It may be that the recognition of rewards as controlling only emerges with age. Children of 5- and 6- tend to have an undifferentiated concept of ability (i.e., ability and effort are perceived as the same; Nicholls, 1984, 1989), therefore, if rewards are offered contingent upon participation then children may attribute rewards for effort put in, rather than competence level.

Amotivation was unexpectedly negatively associated with the controlled types of regulation (external and introjection). This may be because that in this study amotivation was only chosen by two children within the sample. The rest of the children were given a code of 1 (not picked) and coded higher for any other type of regulation they chose, automatically resulting in a negative association. Further research is needed to explore the amotivation aspect as it could provide valuable information for researchers and teachers on how to best support young children.

Taken together, this initial exploration into construct validity indicated that the tool was able to capture some of the hypothesised associations between motivational variables, aligning with past SDT research, though several unexpected relationships were also found. It is possible that the weak relationships among BPNS and types of behavioral regulation observed in this study may change across developmental time and this requires further investigation. Regardless, the correlations observed in the current study should not be over-interpreted due to the small sample size. Future

research should seek to improve MAT-PE before examining construct validity with a larger, more representative sample.

Preliminary descriptives the MAT-PE

Study 3 provided descriptive results from the MAT-PE and codebook that showed promising sensitivity and range in this sample of children. Despite some motivational tools focusing upon intrinsic motivation (Gottfried, 1986, 1990), collapsing introjection and external regulations (Guay et al., 2010), and excluding amotivation, the MAT-PE demonstrates that when given the choice, young children are capable of choosing the types of behavioral regulation underlying their participation in PE. Furthermore, children provided a wide range of need satisfaction responses, supporting the potential of the tool in capturing high and low levels of basic psychological needs.

Enjoyment

This study found that most children perceived PE to be enjoyable. Of the limited studies available for comparison, Baron and Downey (2007) found that PE enjoyment among 7-11-year-olds was high for games, gymnastics and dance activities. PE has been rated as a top 3 favourite subjects in 78% of children aged five to 12 (Coulter & Woods, 2011), which indicates high levels of enjoyment. However, PE enjoyment has been found to decline from the age of nine (Cairney et al., 2012; Prochaska et al., 2003). Therefore, is worth monitoring, especially as it has the potential to predict actual PA and PA intention (Best et al., 2017; Bungum et al., 2000). Our findings are generally in line with previous literature and suggest that the MAT-PE could be used by researchers to identify what young children like and do not like about PE so as to best support overall PE enjoyment.

Basic Psychological Need Satisfaction

Overall, relatedness need satisfaction was high, which aligns with previous literature in older age groups (13-14-year-olds, Ntoumanis et al., 2009; 11-16-year-olds, Taylor et al., 2010). The relatedness items had a possible scoring range of one to four. Children's actual scores for the PE teacher-related items ranged from three to four, while scores on the peer-related items ranged from one to four. This wider range for peer-related items indicates that some children in this sample

chose the more negative options provided to them, indicating sensitivity of the tool and a lack of positive bias on behalf of the children. Although the PE teacher-related items were matched more highly by the independent researchers, we believe the inclusion of both social agents within the tool has the potential to provide useful information. As stated above, this belief is supported by Vasconcellos et al. (2019) who outlined that PE teachers and peers have differential effects upon children's relatedness where peers have more of an effect on relatedness than PE teachers.

Overall, there was lower reported autonomy need satisfaction in comparison to relatedness and competence needs satisfaction in this sample. The possible range for overall autonomy was four to 15 and the actual range was found to be seven to 15. All items demonstrated sensitivity as the actual range was close to the possible range. This more moderate level of autonomy echoes the literature based in older age groups (13-15-year-olds, Taylor & Lonsdale, 2010; nine- to 12-year-olds, van Aart et al., 2017).

Consistent with previous research which examined perceived motor competence in 4-7-year-old children (Noordstar et al., 2016; Spessato et al., 2013), competence need satisfaction was seen to be high within this sample. However, the possible and actual ranges of the global item were similar, indicating some sensitivity in this item among the sample.

Behavioral regulation

Particular sensitivity can be seen in the behavioral regulations aspect of the MAT-PE where the possible and actual ranges were the same, except for external avoidance, which was seen to have an actual range of one to four rather than the possible one to five. This indicates that some children within the sample either did not pick each type of regulation or gave varying responses. Most children were able to provide 'deep' level responses, indicating their comprehension of the items. There was a higher level of autonomous motivation in comparison to controlled, which has been previously found in younger children comparatively to older children (Corpus et al., 2009). However, controlled motivation, or more specifically, external approach was highly present in the

sample. This may be because the use of rewards is considered prevalent within the education system (Deci et al., 2001).

Although only two children in our sample chose amotivation, they both provided deep level responses to the follow-up question indicating that with larger samples it could be further explored what forms of amotivation young children demonstrate in PE. Within a sample of 390 14- to 15-year-olds, only 21 (15 girls, 6 boys) were identified as being amotivated within PE (Ntoumanis et al., 2004). This suggests that the prevalence of amotivation is relatively low in younger and older samples of children and adolescents. We would advocate the inclusion of the amotivation in future versions of the MAT-PE.

Practical Implications and Future Research Directions

As this is a first step towards developing a tool that can assess young children's motivational perceptions in PE, improvements to the tool are needed before more data collection is conducted in a larger, representative sample. The MAT-PE has been shown to be feasible for a researcher to administer one to one with a young child in a quiet location, and the resources are relatively low cost. In all, with 30 minutes allocated to the draw and write classroom-based activity, ~20 minutes for the remainder of the MAT-PE administration, ~60 minutes for transcription and ~30 minutes to code, this equates to around 2 hours per child. Purely quantitative motivation measures have been seen to take from 20-30 minutes to administer (Gottfried, 1990; Guay et al, 2010) on a whole-class basis, while one to one measures, such as The Pictorial Scale of Perceived Movement Skill Competence for Young Children (Barnett et al, 2015) and the Self-Perception Profile for Children (Harter & Pike, 1984) designed for younger children takes less than 10 minutes to administer. Although the MAT-PE has an arguably lengthy administration process, which could be considered a limitation by some, it is worthwhile when considering the amount of depth and richness of data provided by young children. Nevertheless, future research should examine strategies to reduce coding times such as conducting live coding alongside the MAT-PE administration and/or directly coding from audio recordings without transcription. Removing the Write and Draw enjoyment

activity and undertaking live coding could reduce the total administration time to around 20 minutes.

Also, although the tool itself is considered mixed-method, the data it produces can be analysed qualitatively or quantitatively, making it accessible to different types of researchers and research questions. Future studies should also include children with different language and special educational needs (e.g., children with Autism) to assess its accessibility. Future research should also examine further aspects of validity such as predictive validity, as well as test-retest reliability and responsiveness. At this point the MAT-PE is primarily for researcher use, towards understanding young children's motivation, how and if it changes over time, the consequences of motivation in PE on other outcomes such as PA, and to inform interventions. Although it is intended for researcher use, its use within applied research and collaborative Higher Education/elementary education partnerships has relevance for informing teaching practice. This is important as PE teachers can identify children with poor quality or no motivation and their source of motivation and subsequently understand how to support their motivation through their own teaching styles.

Strengths and limitations

This study had several strengths including the comprehensive iterative development of the MAT-PE and codebook with the relevant target populations over a period of 14 months. Strength was also found in the variety of expertise within the research team, where content validity was judged based on multi-disciplinary rather than narrow perspectives (Terwee et al., 2018). A further strength was that independent researchers with a range of SDT experience were sampled in determining the MAT-PE and codebook's acceptability and content validity, enabling the codebook's accessibility to be assessed. Limitations included the involvement of the same children in Study 1 and Study 3. This was due to the present study being conducted within a larger research project (Rudd et al., 2020) with a convenience sample of children that could be accessed within the study timeline. This may have influenced how the children in Study 3 interacted with the tool as it was familiar to them. A further limitation may be attributed to the small number of questions for each

SDT construct due to the young age of the target population, which may restrict the comprehensiveness of the assessment.

Conclusion

This study was a first step towards a novel, mixed-method tool to measure young children's BPNS and behavioral regulations in PE through an age-appropriate set of activities aligned with SDT and informed by young children. The MAT-PE has promising content validity (Study 1), though further development is needed, namely within the autonomy activity (e.g., language of items, addition of volition items), and consideration of the introjected regulation item (i.e., is the ego aspect of introjection better for investigation in this young age group in comparison to feelings of guilt/shame?). The codebook (Study 2) was found to not only have content validity but it was also found to be acceptable and demonstrated excellent reliability. The tool demonstrated some sensitivity, and provided expected and unexpected associations between motivational constructs (Study 3), which requires further investigation. With further development and a larger sample, this tool has the potential to allow researchers to explore how the PE environment affects young children's BPNS and subsequent behavioral regulation. Knowing this information can inform interventions on a class level (changing the environment to support BPN) and to identify individual children who may be experiencing controlled motivation or amotivation. Through this, research can help inform teachers motivating styles and their practice within early primary PE.

Acknowledgements

The authors thank the schools, teachers and children for participating in this research. The authors also thank Laura O'Callaghan, Bethan Price and Adam Lussey for their help with data collection, Martin Cooper for his help with content validity and acceptability of the codebook, Laura Swettenham for her help with content validity, acceptability and reliability of the codebook, and also Dr Amy Whitehead and Dr Gus Ryrie for their help with content validity and acceptability of the codebook. The authors would also like to thank Dr Rob Noonan for his help in developing the tool and Amy Fitton Davies for designing the tool resources.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Bailey, R. (2006). Physical education and sport in schools: A review of benefits and outcomes. *Journal of School Health*, *76*, 397-401. https://doi.org/10.1111/j.1746-1561.2006.00132.x

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*, 98-102. https://doi.org/10.1016/j.jsams.2013.12.004

Baron, L. J., & Downey, P. J. (2007). Perceived success and enjoyment in elementary physical education. *Journal of Applied Research on Learning, 1*, 1-24.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*, 1641-1660.

Bryman, A. (2006). Integrating quantitative and qualitative research: how is it done? *Qualitative research*, *6*, 97-113.

Bungum, T., Dowda, M., Weston, A., Trost, S. G., & Pate, R. R. (2000). Correlates of physical activity in male and female youth. *Pediatric Exercise Science*, *12*, 71-79.

Butler, R. (2005). Competence assessment, competence, and motivation between early and middle childhood. *Handbook of competence and motivation*, 202-221.

Cairney, J., Kwan, M. Y., Velduizen, S., Hay, J., Bray, S. R., & Faught, B. E. (2012). Gender, perceived competence and the enjoyment of physical education in children: a longitudinal examination. *International Journal of Behavioral Nutrition and Physical Activity, 9*, 26.

Carroll, B., & Loumidis, J. (2001). Children's perceived competence and enjoyment in physical education and physical activity outside school. *European Physical Education Review, 7*, 24-43. https://doi.org/10.1177/1356336X010071005

Caruth, G. D. (2013). Demystifying mixed methods research design: A review of the literature. *Online Submission*, *3*, 112-122. https://doi.org/10.13054/mije.13.35.3.2

Casey, A., & Goodyear, V. A. (2015). Can cooperative learning achieve the four learning outcomes of physical education? A review of literature. *Quest, 67*, 56-72. https://doi.org/10.1080/00336297.2014.984733

Chanal, J., Cheval, B., Courvoisier, D. S., & Paumier, D. (2019). Developmental relations between motivation types and physical activity in elementary school children. *Psychology of Sport and Exercise*, *43*, 233-242. https://doi.org/10.1016/j.psychsport.2019.03.006

Chandler, C. L., & Connell, J. P. (1987). Children's intrinsic, extrinsic and internalized motivation: A developmental study of children's reasons for liked and disliked behaviours. *British Journal of Developmental Psychology, 5,* 357-365. https://doi.org/10.1111/j.2044-835x.1987.tb01072.x

Chen, W. (2014). Psychological needs satisfaction, motivational regulations and physical activity intention among elementary school students. *Educational Psychology*, *34*, 495-511. https://doi.org/10.1080/01443410.2013.822959

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290. https://doi.org/10.1037/1040-3590.6.4.284

Coulter, M., & Woods, C. B. (2011). An exploration of children's perceptions and enjoyment of school-based physical activity and physical education. *Journal of Physical Activity and Health*, *8*, 645-654.

Creswell, J. W., & Plano Clark, V. L. (2011). Choosing a mixed methods design. Designing and conducting mixed methods research, 2, 53-106.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 227-268. https://doi.org/10.1207/s15327965pli1104_01

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research, 71*, 1-27. https://doi.org/10.3102/00346543071001001

Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality.

De Meester, A., Van Duyse, F., Aelterman, N., De Muynck, G. J., & Haerens, L. (2020). An experimental, video-based investigation into the motivating impact of choice and positive feedback among students with different motor competence levels. *Physical Education and Sport Pedagogy*, *25*, 361-378.

Department of Education. (2014). The national curriculum in England Framework document [PDF].. Retrieved https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/381344/Master_final_national_curriculum_28_Nov.pdf

Dunn, J. G., Bouffard, M., & Rogers, W. T. (1999). Assessing item content-relevance in sport psychology scale-construction research: Issues and recommendations. *Measurement in Physical Education and Exercise Science*, *3*, 15-36.

Erickson, J. D. (2019). Primary Students' Emic Views of Reading Intervention: A Qualitative Case Study of Motivation. *Literacy Research: Theory, Method, and Practice, 68*, 86-107. https://doi.org/10.1177/2381336919870282

Erwin, H. E., Stellino, M. B., Beets, M. W., Beighle, A., & Johnson, C. E. (2013). Physical education lesson content and teacher style and elementary students' motivation and physical activity levels. *Journal of Teaching in Physical Education, 32,* 321-334. https://doi.org/10.1123/jtpe.32.3.321

Escriva-Boulley, G., Tessier, D., Ntoumanis, N., & Sarrazin, P. (2018). Need-supportive professional development in elementary school physical education: Effects of a cluster-randomized control trial on teachers' motivating style and student physical activity. *Sport, Exercise, and Performance Psychology, 7*, 218-234. https://doi.org/10.1037/spy0000119

Evans, P., & Fuller, M. (1996). 'Hello. Who Am I Speaking To?' Communicating with Pre-school Children in Educational Research Settings. Early Years, 17(1), 17-20.

Evans, P., & Fuller, M. (1998). Children's perceptions of their nursery education. International Journal of Early Years Education, 6(1), 59-74.

Fernald, D. H., & Duclos, C. W. (2005). Enhance your team-based qualitative research. *The Annals of Family Medicine, 3*, 360-364. https://doi.org/10.1370/afm.290

Ferrer-Caja, E., & Weiss, M. R. (2000). Predictors of intrinsic motivation among adolescent students in physical education. *Research quarterly for exercise and sport*, *71*, 267-279.

Fonteyn, M. E., Vettese, M., Lancaster, D. R., & Bauer-Wu, S. (2008). Developing a codebook to guide content analysis of expressive writing transcripts. *Applied Nursing Research, 21*, 165-168. https://doi.org/10.1016/j.apnr.2006.08.005

Gao, Z., Podlog, L., & Huang, C. (2013). Associations among children's situational motivation, physical activity participation, and enjoyment in an active dance video game. *Journal of Sport and Health Science*, *2*, 122-128.

Gelman, R., & Baillargeon, R. (1983). Review of some Piagetian concepts. *Handbook of child psychology*: formerly Carmichael's Manual of child psychology/Paul H. Mussen, editor.

Gottfried, A. E. (1986). *CAIMI: Children's academic intrinsic motivation inventory*. Psychological Assessment Resources.

Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology, 82,* 525-538. https://doi.org/10.1037/0022-0663.82.3.525

Guay, F., Chanal, J., Ratelle, C. F., Marsh, H. W., Larose, S., & Boivin, M. (2010). Intrinsic, identified, and controlled types of motivation for school subjects in young elementary school children. *British Journal of Educational Psychology*, *80*, 711-735. https://doi.org/10.1348/000709910x499084

Harter, S. (1988). Issues in the assessment of the self-concept of children and adolescents. In A. LaGreca (Ed.), *Childhood assessment: Through the eyes of a child* (pp. 292–325). Allyn & Bacon.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child development*, *55*, 1969-1982. https://doi.org/10.2307/1129772

Hills, A. P., Dengel, D. R., & Lubans, D. R. (2015). Supporting public health priorities: recommendations for physical education and physical activity promotion in schools. *Progress in Cardiovascular Diseases, 57*, 368-374. ttps://doi.org/10.1016/j.pcad.2014.09.010

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Thomson.

Huhtiniemi, M., Sääkslahti, A., Watt, A., & Jaakkola, T. (2019). Associations among basic psychological needs, motivation and enjoyment within finnish physical education students. *Journal of Sports Science & Medicine*, *18*, 239.

Irwin, L. G., & Johnson, J. (2005). Interviewing young children: Explicating our practices and dilemmas. *Qualitative Health Research, 15*, 821-831. https://doi.org/10.1177/1049732304273862

Jaakkola, T., Washington, T., & Yli-Piipari, S. (2013). The association between motivation in school physical education and self-reported physical activity during Finnish junior high school: A self-determination theory approach. *European Physical Education Review*, *19*, 127-141. https://doi.org/10.1177/1356336x12465514

Katz, I., Kaplan, A., & Buzukashvily, T. (2011). The role of parents' motivation in students' autonomous motivation for doing homework. *Learning and Individual differences, 21*, 376-386. https://doi.org/10.1016/j.lindif.2011.04.001

Kimiecik, J. C., & Harris, A. T. (1996). What is enjoyment? A conceptual/definitional analysis with implications for sport and exercise psychology. *Journal of Sport and Exercise Psychology, 18*, 247-263.

Kirk, D. (2005). Physical education, youth sport and lifelong participation: the importance of early learning experiences. *European Physical Education Review*, *11*, 239-255.

Ladwig, M. A., Vazou, S., & Ekkekakis, P. (2018). “My best memory is when I was done with it”: PE memories are associated with adult sedentary behavior. *Translational Journal of the American College of Sports Medicine, 3*, 119-129.

Leptokaridou, E. T., Vlachopoulos, S. P., & Papaioannou, A. G. (2016). Experimental longitudinal test of the influence of autonomy-supportive teaching on motivation for participation in elementary school physical education. *Educational Psychology, 36*, 1138-1159. https://doi.org/10.1080/01443410.2014.950195

MacQueen, K. M., McLellan, E., Kay, K., & Milstein, B. (1998). Codebook development for team-based qualitative analysis. *Cam Journal, 10*, 31-36. https://doi.org/10.1177/1525822x980100020301

Malti, T. (2016). Toward an integrated clinical-developmental model of guilt. *Developmental Review*, *39*, 16-36.

Markland, D., & Tobin, V. (2004). A modification to the behavioural regulation in exercise questionnaire to include an assessment of amotivation. *Journal of Sport and Exercise Psychology*, *26*, 191-196.

Marques, A., Corrales, F. R. G., Martins, J., Catunda, R., & Sarmento, H. (2017). Association between physical education, school-based physical activity, and academic performance: a systematic review. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, *31*, 316-320.

Mellor, D., & Moore, K. A. (2014). The use of Likert scales with children. *Journal of Pediatric Psychology, 39*, 369-379. https://doi.org/10.1093/jpepsy/jst079

Milyavskaya, M., & Koestner, R. (2011). Psychological needs, motivation, and well-being: A test of self-determination theory across multiple domains. *Personality and Individual Differences, 50*, 387-391. https://doi.org/10.1016/j.paid.2010.10.029

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., ... & De Vet, H. C. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Medical Research Methodology*, *10*, 1-8.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & de Vet, H. C. (2012). COSMIN checklist manual. *Amsterdam: University Medical Center*.

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91*, 328.

Nicholls, J. G. (1989). *The competitive ethos and democratic education*. Harvard University Press. doi: https://doi.org/10.5860/choice.27-1049

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,* 326. https://doi.org/10.1186/s12889-016-3005-1

Noordstar, J. J., van der Net, J., Jak, S., Helders, P. J., & Jongmans, M. J. (2016). The change in perceived motor competence and motor task values during elementary school: A

longitudinal cohort study. *British Journal of Developmental Psychology, 34*, 427-446. https://doi.org/10.1111/bjdp.12142

Ntoumanis, N. (2001). A self-determination approach to the understanding of motivation in physical education. *British Journal of Educational Psychology*, *71*, 225-242.

Ntoumanis, N. (2005). A prospective study of participation in optional school physical education using a self-determination theory framework. *Journal of Educational Psychology*, *97*, 444-453. https://doi.org/10.1037/0022-0663.97.3.444

Ntoumanis, N., Pensgaard, A. M., Martin, C., & Pipe, K. (2004). An idiographic analysis of amotivation in compulsory school physical education. *Journal of Sport and Exercise Psychology, 26*, 197-214. https://doi.org/10.1123/jsep.26.2.197

Ntoumanis, N., Barkoukis, V., & Thøgersen-Ntoumani, C. (2009). Developmental trajectories of motivation in physical education: Course, demographic differences, and antecedents. *Journal of Educational Psychology*, *101*, 717.

Ponce, O. A., & Pagán-Maldonado, N. (2015). Mixed methods research in education: Capturing the complexity of the profession. *International Journal of Educational Excellence, 1*, 111-135. https://doi.org/10.18562/ijee.2015.0005

Porcellato, L., Dugdill, L., & Springett, J. (2005). A longitudinal study exploring Liverpool primary schoolchildren's perspectives on smoking. *Childhood, 12,* 425-443. https://doi.org/10.1177/0907568205058601

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*, 1147-1157. https://doi.org/10.1007/s11136-018-1798-3

Prinsen, C. A., Vohra, S., Rose, M. R., Boers, M., Tugwell, P., Clarke, M., ... & Terwee, C. B. (2016). How to select outcome measurement instruments for outcomes included in a "Core Outcome Set"–a practical guideline. *Trials*, *17*, 1-10.

Prochaska, J. J., Sallis, J. F., Slymen, D. J., & McKenzie, T. L. (2003). A longitudinal study of children's enjoyment of physical education. *Pediatric Exercise Science, 15*, 170-178.

Reilly, J. J. (2016). When does it all go wrong? Longitudinal studies of changes in moderate-to-vigorous-intensity physical activity across childhood and adolescence. *Journal of Exercise Science & Fitness, 14*, 1-6. https://doi.org/10.1016/j.jesf.2016.05.002

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational technology research and development*, *52*, 5.

Rudd, J. R., Crotti, M., Fitton-Davies, K., O'Callaghan, L., Bardid, F., Utesch, T., ... & Foweather, L. (2020). Skill Acquisition Methods Fostering Physical Literacy in Early-Physical Education (SAMPLE-PE): Rationale and study protocol for a cluster randomized controlled trial in 5–6-year-old children from deprived areas of North West England. *Frontiers in Psychology*, *11*, 1228.

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, *57*, 749.

Ryan, R. M., & Deci, E. L. (2017). *Self—Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness.* Guildford Press.

Ryan, R. M., & Patrick, H. (2009). Self-determination theory and physical activity: the dynamics of motivation in development and wellness. *Hellenic Journal of Psychology, 6*, 107-124.

Sallis, J. F., & McKenzie, T. L. (1991). Physical education's role in public health. *Research Quarterly for Exercise and Sport, 62*, 124-137. https://doi.org/10.1080/02701367.1991.10608701

Sallis, J. F., McKenzie, T. L., Beets, M. W., Beighle, A., Erwin, H., & Lee, S. (2012). Physical education's role in public health: Steps forward and backward over 20 years and HOPE for the future. *Research Quarterly for Exercise and Sport, 83,* 125-135. https://doi.org/10.5641/027013612800745329

Sebire, S. J., Jago, R., Fox, K. R., Edwards, M. J., & Thompson, J. L. (2013). Testing a self-determination theory model of children's physical activity motivation: a cross-sectional study. International

Journal of Behavioural Nutrition and Physical Activity, 10, 111. https://doi.org/10.1186/1479-5868-10-111

Smith, N., Tessier, D., Tzioumakis, Y., Quested, E., Appleton, P., Sarrazin, P., ... & Duda, J. L. (2015). Development and validation of the multidimensional motivational climate observation system. *Journal of sport and exercise psychology*, *37*(1), 4-22.

Solmon, M. A., & Carter, J. A. (1995). Kindergarten and first-grade students' perceptions of physical education in one teacher's classes. *The Elementary School Journal*, *95*, 355-365.

Spessato, B. C., Gabbard, C., Robinson, L., & Valentini, N. C. (2013). Body mass index, perceived and actual physical competence: the relationship among young children. *Child: Care, Health and Development, 39*, 845-850. https://doi.org/10.1111/cch.12014

Standage, M., Duda, J. L., & Ntoumanis, N. (2003). A model of contextual motivation in physical education: Using constructs from self-determination and achievement goal theories to predict physical activity intentions. *Journal of Educational Psychology*, *95*, 97-110. https://doi.org/10.1037/0022-0663.95.1.97

Standage, M., Duda, J. L., & Ntoumanis, N. (2005). A test of self-determination theory in school physical education. *British Journal of Educational Psychology, 75,* 411-433. https://doi.org/10.1348/000709904x22359

Standage, M., Gillison, F. B., Ntoumanis, N., & Treasure, D. C. (2012). Predicting students' physical activity and health-related well-being: A prospective cross-domain investigation of motivation across school physical education and exercise settings. *Journal of Sport and Exercise Psychology, 34*, 37-60. https://doi.org/10.1123/jsep.34.1.37

Stefanou, C. R., Perencevich, K. C., Dicintio, M., & Turner,J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist, 39*, 97–110. https://doi.org/10.1207/s15326985ep3902_2

Stone, W. L., & Lemanek, K. L. (1990). Developmental issues in children's self-reports. *Journal of Autism and Developmental Disorders, 20,* 513-522. https://doi.org/10.1007/bf02216056

Taylor, I. M., & Lonsdale, C. (2010). Cultural differences in the relationships among autonomy support, psychological need satisfaction, subjective vitality, and effort in British and Chinese physical education. *Journal of Sport and Exercise Psychology*, *32*, 655-673.

Taylor, I. M., & Ntoumanis, N. (2007). Teacher motivational strategies and student self-determination in physical education. *Journal of educational psychology*, *99*, 747-760. https://doi.org/10.1037/0022-0663.99.4.747

Taylor, I. M., Ntoumanis, N., Standage, M., & Spray, C. M. (2010). Motivational predictors of physical education students' effort, exercise intentions, and leisure-time physical activity: A multilevel linear growth analysis. *Journal of Sport and Exercise Psychology*, *32*, 99-120.

Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: a systematic review. *International journal of behavioral nutrition and physical activity*, *9*(1), 78.

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*, 1159-1170. https://doi.org/10.1007/s11136-018-1829-0

Tsangaridou, N., & Lefteratos, C. (2013). Elementary students' views and experiences on Sport Education in Cyprus. *Advances in Physical Education, 3*, 28-35. https://doi.org/10.4236/ape.2013.31005

United Nations Educational Scientific and Cultural Organisation. (2013). *World-wide survey of school physical education*. Retrieved https://bit.ly/35QveBq

Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. *Advances in Experimental Social Psychology*, *29*, 271-360.

Vallerand, R. J. (2007). Intrinsic and extrinsic motivation in sport and physical activity: A review and a look at the future.

Vallerand, R. J., Blais, M. R., Brière, N. M., & Pelletier, L. G. (1989). Construction et validation de l'échelle de motivation en éducation (EME). *Canadian Journal of Behavioural Science, 21*, 323-349. https://doi.org/10.1037/h0079855

van Aart, I., Hartman, E., Elferink-Gemser, M., Mombarg, R., & Visscher, C. (2017). Relations among basic psychological needs, PE-motivation and fundamental movement skills in 9–12-year-old boys and girls in Physical Education. *Physical Education and Sport Pedagogy*, *22*, 15-34.

Van der Kaap-Deeder, J., Vansteenkiste, M., Soenens, S., Loeys, T., Mabbe, E., & Gargurevich, R. (2015). Autonomy-supportive parenting and autonomy-supportive sibling interactions: The role of mothers' and siblings' psychological need satisfaction. *Personality and Social Psychology Bulletin, 41*, 1590-1604.

Vansteenkiste, M., Zhou, M., Lens, W., & Soenens, B. (2005). Experiences of autonomy and control among Chinese learners: Vitalizing or immobilizing? Journal of educational psychology, 97(3), 468.

Vasconcellos, D., Parker, P. D., Hilland, T., Cinelli, R., Owen, K. B., Kapsal, N., Antczak, D., Ntoumanis, N., Ryan, R. M., & Lonsdale, C. (2019). Self-determination theory applied to physical education: A systematic review and meta-analysis. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000420

White, M. D., & Marsh, E. E. (2006). Content analysis: A flexible methodology. *Library Trends*, *55*, 22-45. https://doi.org/10.1353/lib.2006.0053

Whitehead, M. (Ed.). (2019). *Physical Literacy Across the World*. Routledge. https://doi.org/10.4324/9780203702697

Wiering, B., de Boer, D., & Delnoij, D. (2017). Patient involvement in the development of patient-reported outcome measures: a scoping review. *Health Expectations*, *20*(1), 11-23.

Woods, S. E., Springett, J., Porcellato, L., & Dugdill, L. (2005). 'Stop it, it's bad for you and me': experiences of and views on passive smoking among primary-school children in Liverpool. *Health Education Research, 20*, 645-655. https://doi.org/10.1093/her/cyh02

Yli-Piipari, S., Watt, A., Jaakkola, T., Liukkonen, J., & Nurmi, J. E. (2009). Relationships between physical education students' motivational profiles, enjoyment, state anxiety, and self-reported physical activity. *Journal of Sports Science & Pedicine, 8*, 327.

Highlights

- Developed the Motivation Assessment Tool for Physical Education (MAT-PE)
- Assesses basic psychological need satisfaction and behavioural regulation in PE
- Mixed method assessment
- Suitable for use with young children (5-6 years old)
- Initial findings demonstrate that the MAT-PE has promising validity
- More research is required to further develop and improve the MAT-PE

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: