

Motivation Scale Exploratory Development in Chinese University Tennis Physical Education Context

SAGE Open
Vol. 16(3): 1–22
© The Author(s) 2026
DOI: 10.1177/21582440261479388
journals.sagepub.com/home/sgo
 Sage

Changjie Liu^{1,2} , Nung Kion Lee³ , and Kee-Man Chuah⁴

Abstract

Objective: The present research aimed to evaluate the psychometrics of Perceived Locus of Causality (PLOC) in the Chinese compulsory tennis physical education context. **Method:** Two Samples (n1=355, n2=245) of data were obtained from one Chinese university students who had tennis PE in semester 2024/2025-1 and 2024/2025-2. In Phase 1, the Self-determination Theory (SDT) based 5-factor model was examined for the total data. Based on Phase 1 implications, the original 5-factor model was shortened into a 4-factor model in Phase 2 and a 3-factor model in Phase 3. In Phase 4, a two-procedure EFA-CFA was performed across the two independent natural samples. Post-hoc Model Modification strategies were applied for the 4-phase analysis, including: (1) building paths stepwise based on values of modification indices (MI) and expected parameter change (EPC); (2) detecting low-loading items in CFA, (3) checking cross-loading items in EFA. **Results:** The modified five-factor model obtained a good fit, but not discriminant validity. An SDT-based 3-factor model and an AGT-SDT-based 3-factor model were developed with good reliability and validity for the present population. **Conclusion:** A 17-item 1-path 3-factor model obtained acceptable indices and converged with the simplified SDT continuum of Autonomous Motivation, Controlled Motivation, and Amotivation. The EFA-CFA analysis procedure derived an 18-item 3-factor model (Approach Motivation, Avoidance Motivation, and Amotivation), performed with acceptable quality and developmental adaptations in the present population.

Keywords

PLOC, EFA, CFA, Self-determination Theory, achievement goal theory

Received: 30 September 2025; Revised: 18 January 2026; Accepted: 5 August 2026

Introduction

Physical education (PE) plays a vital role in promoting students' physical health, mental well-being (Muratbayevich, 2024), and social development (Sallis et al., 2012; Salmanov et al., 2025; Wuest & Bucher, 2018). Thus, it is critical to understand students' motivation for enhancing PE at school. Achievement Goal Theory (AGT) and Self-determination Theory (SDT) are two prominent theories that guided investigations on motivational behavior in the past decades (Chen & Solmon, 2024).

¹Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak, Kota Samarahan, Malaysia

²School of Physical Education, Weifang Institute of Technology, Weifang, Shandong, China

³Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, Kota Samarahan, Malaysia

⁴Faculty of Education, Language and Communication, Universiti Malaysia Sarawak, Kota Samarahan, Malaysia

Corresponding author:

Nung Kion Lee, Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak 94300, Malaysia.

Email: nklee@unimas.my



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

PLOC

SDT explains human motivation by emphasizing people's inherent tendencies toward growth and the role of innate psychological needs in driving self-motivated behavior (Deci & Ryan, 1985). Perceived Locus of Causality (PLOC) is a foundational concept in SDT, which refers to an individual's subjective perceptions of the causes of their behaviors (Ryan & Connell, 1989). PLOC categorizes these causes (motivation) into intrinsic motivation (IM), identified regulation (IDR), introjected regulation (ITR), external regulation (ER), and amotivation (AM). IM is about inherent satisfaction engaging in activities with spontaneous feelings of effectance, interest, or enjoyment (Ryan & Deci, 2017), such as enjoying learning new skills like sport techniques, and finding the activity exciting. IDR portrays that people are not simply complying with an external or introjected demand but are instead acting out of a belief in the personal importance or perceived value of the activity (Ryan & Deci, 2017), for example, wanting to learn sport skills to improve fitness and health, and realizing the usefulness of obtaining swimming skills. IDR behaviors reflect a high degree of volition, where the regulation is internalized and experienced as self-congruent, sharing qualities of autonomy with intrinsic motivation. Therefore, IM and IDR were defined as autonomous motivation in a simplified motivational continuum (Deci & Ryan, 2008).

ITR is under less internal pressure, tension, and conflict; individuals avoid shame or acquire external/internal rewards like praise and pride (Ntoumanis et al., 2009), as values are incompletely assimilated into individuals' beliefs (Ryan & Deci, 2017). Behaviors of ITR in PE are like wanting teachers/peers to view them as good/skillful, or avoiding feeling bad about non-participation in PE class or group learning activities. ER defines individuals who avoid punishment, try to obtain praise and reward (Ntoumanis et al., 2009). Examples include attending class to avoid getting into trouble and following rules to earn a grade. AM defines that individuals who lack aim or reasons to do something, and don't have the intention to achieve an outcome. In physical education, Amotivated students often perceive the class as pointless or a waste of time. This leads to observable disengagement, such as sitting out during games, showing clear disinterest, or even quitting activities, often because they feel incompetent.

PLOC was initially developed by Goudas et al. (1994), adopted from Ryan and Connell (1989), and measured ER, ITR, IDR, and IM for exercise. Then, Vallerand et al. (1992) expanded the motivation spectrum by adding AM. Wolf et al. (2022) refined the scale (Goudas et al., 1994) by including integrated regulation and validating the new six-factor PLOC. Furthermore, the equivalence of PLOC was examined through language (Wolf et al., 2022), gender (Abós et al., 2018; Teixeira et al., 2018), and culture (Wang et al., 2009; Yang et al., 2019). Teixeira et al. (2018) translated the five-factor model into the Portuguese context to ensure proper psychometric validation in both gender and calibration/validation samples. Yang et al. (2019) tested psychometric properties equivalence between Chinese (n=1287) and Spanish (n=1006) samples, and found that Chinese individuals scored higher on intrinsic motivation and identified motivation, but with lower introjected regulation, external regulation, and amotivation. Besides, Wang et al. (2009) identified that there were variations in the mean levels of PLOC components across cultures, as the Singapore sample had higher scores than the British sample.

However, some researchers found that their findings did not fully meet psychometric requirements. Lonsdale et al. (2011) indicated that introjected and external regulation lacked internal consistency in the Hong Kong Chinese sample, as well as that the validity of the distinction was not sufficient, as intrinsic motivation was highly correlated with identified motivation. This might be caused by contingent self-worth in contrast to existing motives (Vlachopoulos et al., 2011).

In conclusion, PLOC has been widely applied in motivation research, though it has not yet been robustly validated. First, the item reliability on ITR and ER was often reported with unacceptable values. Second, intrinsic motivation (IM) had a high correlation with identified regulation, and the two constructs were not robustly discriminant. Third, the dissatisfaction with the fulfillment of consistency and divergent validity of PLOC greatly influence relevant studies. Therefore, dependable strategies are essential for connecting PLOC validation with subsequent construction selection. Lastly, scale validation across different cultures, languages, and sports is essential for PLOC development.

AGT and SDT Converge

A relative motivation theory, AGT, standing on a micro perspective, focusing on task achievement, assumes that the goal of humans is the drive for individuals to achieve competence (Nicholls, 1984). While SDT, as a macro theory, focuses on social-contextual factors that support or thwart an individual's basic psychological needs (BPN) in competence, relatedness, and autonomy, and places them in biological underpinnings and an evolutionary perspective (Ryan & Deci, 2017).

In fact, existing studies have indicated the connections and integrations between AGT and SDT from the perspectives of 1) convergence in defining the learning environment, 2) using them simultaneously for describing learning behavior by disclosing associations of constructs from the two theories, and 3) linking constructs from the two theories to produce integrative theoretical frameworks.

Firstly, the two theories share common concepts. For example, SDT sorts the educational environment into BPN-supportive or BPN-thwarted climate, while the learning environment in AGT is defined as mastery- or performance-oriented goal (Ames, 1992). Both theories propose that intrinsic motivation is nurtured in autonomous and self-mastery situations, but thwarted in social comparison, normative comparison, and performance-rewarding climate (Standage et al., 2003).

Secondly, constructs from the two theories were examined to have associations. The two types of climates, named as task- and ego directions, were examined to be reciprocally associated with BPN satisfaction (García-González et al., 2025). BPN is a preliminary concept in SDT for predicting motivation levels. In another study, path analysis demonstrated that mastery/performance-avoidance goals mediated the relationship between autonomous motivation/introjected-external regulation and educational outcomes/engagement (Rothes et al., 2022).

Thirdly, theoretical frameworks are expanded and updated based on integrations of the two theories. For instance, the construct of mastery and ego-centric climate from AGT, and autonomy-social supportive and controlling environment from SDT were integrated into empowering and disempowering climate (Duda & Appleton, 2016). Urhahne and Wijnia (2023) reorganized the components from AGT, SDT, and other theories to formulate a series of frameworks. Thus, motivational theories are increasingly expanded and differentiated, and fit for debating the changing understanding of motivational components (Schunk et al., 2014; Urhahne & Wijnia, 2023). In addition, Vansteenkiste et al. (2014) proposed to add examinations of reasons underlying achievement goals, already described by SDT, for future prospective development of AGT. These conceptual and empirical evidence supported alternative options for motivation scale development. The converged connections between SDT and AGT support accessibility for coordinating real-world situations and theory explanations.

In scale development practice, there is a common concern that existing theoretical models may not fit real-world phenomena (MacCallum & Tucker, 1991), for instance, the reliability and validity of PLOC have not been robustly examined in previous studies. The present research aimed to gather reliability and validity evidence from two samples to justify the SDT-based model. A series of exploratory procedures was used for reexamining the true model for the present population. The implications for measurement model selection and theory development integration between SDT and AGT are discussed.

Methods

Participants

Sample size for factor analysis was suggested by the guiding rules of thumb, in which Hair et al. suggested 100 or greater (Hair, 1995). A small sample size was identified to increase the risk of Type I error (false positives) and Type II error (false negatives). Wolf et al. (2013) suggested that a minimum sample size of 150 was required for a model, which was 3 factors, 3 or 4 indicators for each factor, and loadings of 0.80, while a model with loadings of 0.65, 3 factors, and 3 or 4 indicators for each factor was suggested to require slightly more than 200 participants for the sample size.

Three hundred and fifty-five participants (Sample 1, 2024-2025-1) and two hundred and forty-five participants (Sample 2, 2024-2025-2) who autonomously registered for the two semesters of tennis PE from one university named Weifang Institute of Technology, in Shandong, China were recruited. Consent was obtained from participants in written form. Participants were informed about the research aim through the distribution of an electronic form, and their responses would be kept confidential and anonymous. They were also informed that they could freely choose to respond or not.

Instrument

The translation and cross-cultural adaptation of the PLOC into the Chinese language for educational application was conducted in accordance with established guidelines for educational and psychological testing (International Test Commission, 2018) and widely adopted adaptation frameworks (Hambleton & Patsula, 1998). The procedure involved a multi-stage design to prioritize construct and functional equivalence within the Chinese University PE context. In line with the International Test Commission guidance and established test-adaptation frameworks, our adaptation decisions prioritized functional and conceptual equivalence over literal word matching. Cultural and semantic refinements were therefore restricted to wording-level adjustments that improved naturalness and interpretability for Chinese university PE students, while preserving the intended regulatory meaning in SDT terms. During the expert committee review, each item was checked against the original construct definition and evaluated for semantic fidelity, pragmatic appropriateness in classroom discourse, and potential shifts in motivational “signal” (autonomy-supportive versus controlling phrasing). Any revisions were agreed through committee consensus, and an audit trail of item-by-item changes and rationales is provided in Table 1 to support transparency and equivalence evaluation. The pre-final version was pilot-tested via cognitive interviews with a

Table 1. Translation and Adaptation of PLOC

Item	PLOC	Items translation and revision	Revision reasons
		我参加体育课是因为... I participate in PE because...	
IM1	Because physical education is fun.	我感觉体育课很有趣。 I feel PE is interesting.	Move “Because” to the front of this scale for concision. Use the first-person pronouns to engage participants’ earnestness in filling out the item.
IM2	Because I enjoy learning new skills.	我喜欢学习新技能。 I enjoy learning new skills.	For euphemistic expression in Chinese culture.
IM3	Because physical education is stimulating.	体育运动让人感觉兴奋开心。 PE makes people feel excited and happy.	
IM4	For the satisfaction I feel while learning new skills/ techniques.	学习运动技能有满足感。 Learning skills brings satisfaction.	Referenced from Vlachopoulos et al. (2011) , and being discriminant from item 2, especially in the Chinese language.
IDR1	Because I want to learn PE/ sports skills.	学习运动技能对我来说很重要。 Learning skills is important for me.	
IDR2	Because it is important for me to do well in physical education.	运动方面有好的表现对我来说很重要。 It is important for me to do well in PE.	A simple expression in Chinese
IDR3	Because I want to improve in sport.	我想在运动方面有所进步。 I want to improve in sports.	
IDR4	Because I can learn skills that I could use in other areas of my life.	我可以在生活其他领域使用体育课学到的运动技能。 I can learn skills that I could use in other areas of my life.	To exemplify a negative situation that usually happens to trigger participants’ responses.
ITR1	Because I want the teacher to think that I am a good student.	我想让老师认为我是一个好学生。 I want the teacher to think that I am a good student.	
ITR2	Because I would feel bad about myself if I didn’t.	我还没掌握运动技能时, 我会担心焦虑。 I would feel anxious when I haven’t mastered a sports skill.	
ITR3	Because I want the other students to think that I am clever.	我想让其他学生认为我很聪明。 I want the other students to think that I am clever.	
ITR4	Because it worries me when I don’t.	若体育课表现不佳, 这会使我担心。 It makes me worried when I don’t perform well.	
ER1	Because I will have problems if I don’t.	不认真练习技能, 期末考核就会遇到麻烦。 I will have problems if I don’t practice the skill seriously.	
ER2	Because that’s what I’m supposed to do.	运动技能练习, 本就是体育课中应该做的事。 Practicing skills is what I’m supposed to do.	

(continued)

Table 1. (continued)

Item	PLOC	Items translation and revision	Revision reasons
		我参加体育课是因为... I participate in PE because...	
ER3	So that the teacher doesn't yell at me.	我认真上体育课, 老师就不会批评我。 If I take PE class seriously, the teacher won't criticize me.	For euphemistic expression in Chinese culture, replace "yell" with "criticize".
ER4	Because that's the rule.	上体育课是学生应当遵循的规则。 That's the rule for students to have a PE lesson.	
AM1	But I don't really know why.	不知道为什么要运动技能练习。 I don't really know why we should practice skills.	Post skill practices to exemplify PE activities.
AM2	But I don't understand why we should have physical education.	不知道为什么我们应该上体育课。 I don't understand why we should have physical education.	
AM3	But I really feel like I'm wasting my time on PE.	体育课很浪费时间。 PE really wastes my time.	
AM4	But I can't see what I'm getting out of PE.	不明白从体育课上得到了什么。 I can't see what I'm getting out of PE.	

purposive sample of 20 students and five teachers to assess clarity, cultural relevance, and perceived difficulty. Insights from this validation phase informed final revisions, yielding the adapted scale ready for full psychometric evaluation of its reliability and validity in the new educational setting.

The PLOC has a clear theoretical core in SDT, but it has also appeared in multiple published implementations and refinements across PE and exercise settings, reflecting ongoing psychometric development rather than a single immutable item set. In the present study, we therefore treated the instrument as a context-adapted PLOC item set for tennis PE rather than as a verbatim reproduction of one historical form. Most items were adapted from the original PE-oriented version (Goudas et al., 1994), and a small number were selected from later PLOC applications to strengthen construct coverage and item clarity in this target context. Importantly, this approach did not introduce new motivational constructs. All items were mapped a priori to the established PLOC regulation types, namely intrinsic motivation, identified regulation, introjected regulation, external regulation, and amotivation.

Then, one item about Amotivation, 'But I can't see what I'm getting out of PE.' was incorporated from Vlachopoulos et al. (2011). And one item about IDR, 'Because I can learn skills that I could use in other areas of my life.' (Lonsdale et al., 2011) was introduced. One item about IM, 'For the satisfaction I feel while learning new skills/techniques.' was translated from (Murcia et al., 2009). Finally, the number of items on each factor was four. The questionnaire was distributed to participants through the online platform Wen Juan Xing before the tennis PE examination. Students of Sample 1 responded with answers on a 5-point Likert scale (1 = not agree at all and 5 = totally agree). Raw data of Sample 2 were 0-100 points. In Phases 1, 2, and 3, the two-sample data were transformed using the z-score method and combined for subsequent CFA. It was unnecessary to transform the two independent datasets for Phase 4 because data analysis was performed separately.

Data Analysis

Prior to formal data analysis, the interquartile range (IQR) method is used for identifying potential outliers within a nonparametric or unknown distribution data set (Tukey, 1977), which is considered to be robust to detect outliers, as it is not influenced by extreme observations (Sullivan et al., 2021).

EFA

And then, exploratory factor analysis (EFA) is applied in R 4.5.1 (psych) to deduce a large number of observed indicators into a smaller number of hidden dimensions or factors (Spearman, 1904; Tinsley & Tinsley, 1987). Five essential techniques of EFA method pave the way to identify underlying constructs as: (1) test appropriateness of factor analysis, (2) factor extraction method selection, (3) number of factors retained by multiple criteria convergence, (4) rotation method selection, and (5) interpretation and construct labeling (DeVellis & Thorpe, 2022; Watkins, 2021).

Before extracting the factors, Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (>0.5 is acceptable) (Kaiser, 1970; Kaiser & Rice, 1974), and Bartlett's Test of Sphericity (p -value <0.05 is good and shows factorability of variables), are performed to evaluate the suitability of observed data for factor analysis (Bartlett, 1950). Then, the Principal Axis Factoring (PAF) extraction method and Promax rotation method are applied to identify the factor structure of translated PLOC. PAF is suitable for non-normal indicators, focused on shared variance among items, when translated scales (cultural adaptation) may influence measurement error (Fabrigar et al., 1999; Suleymanova & Zangieva, 2022).

Next, multiple criteria are simultaneously used to determine factor extraction, including eigenvalue > 1 (Kaiser, 1960), Scree plot (Cattell, 1966), and a cumulative percentage of variance extracted greater than 50% (Horn, 1965). Because researchers reach a consensus that the correct number of factors is almost retrieved when these criteria converge (Auerswald & Moshagen, 2019). Parallel analysis compares the actual eigenvalues from observed data with the average eigenvalues obtained from multiple random datasets (Horn, 1965). The number of factors determined by parallel analysis is when the factor number for the explained variance in observed data is larger than the simulated data (Saqr & López-Pernas, 2024).

After a number of factors are extracted, the Rotation method approaches are used to make data more interpretable without changing its mathematical properties, and more specifically, to maximize high correlations between factors and minimize low ones (Tabachnick et al., 2019). Considering the correlations among factors within PLOC, the oblique (oblimin) method was chosen for the present research.

The pattern matrix is regarded as the primary result for researchers to report when using the oblique method. Because oblique considers correlation among factors, the pattern matrix is the final solution to obtain a simple structure (Fabrigar & Wegener, 2012).

Factor loading is the correlation between the item and the factor (Tavakol & Wetzel, 2020). Factor loading is high if it is more than 0.60, and it is moderately high when it is above 0.30 (Kline, 2014). Notably, a cross-loading item is an item that has a correlation value of more than 0.32 on two or more factors, which explains 10% overlapping variance with other items on that factor (Costello & Osborne, 2005; Tabachnick & Fidell, 2001). It would influence parallel analysis, CFI, and TLI (Li et al., 2020), as well as increasing the critical performance, raising the risk of the type I error rate (Auerswald & Moshagen, 2019). Cross-loading items will be removed. But the number of items on one factor is suggested to be no less than 3 (Fabrigar & Wegener, 2012).

CFA

CFA with MLR calculation and model revision performed in Amos 24.0.0 and R 4.5.1 (lavaan, and semTools). Model revision could iteratively proceed based on modification indices (MI), expected parameter change (EPC), as well as theoretical support and empirical evidence. MI reflects the extent to which a hypothesized model is appropriately described, which captures the evidence of misspecification and is conceptualized as a χ^2 statistic with one degree of freedom (Byrne, 2016). Parameters with MI values greater than 10 were suggested to be freely estimated, expecting a drop in χ^2 in the next round of runs (Byrne, 2016). EPC is the approximated value for estimating a relationship when a parameter is added to a model (Sarlis et al., 1987; Whittaker & Schumacker, 2022).

One parameter could be added if it were associated with large and statistically significant values of MI and EPC, and a theoretical justification (R. C. MacCallum et al., 1992; Whittaker & Schumacker, 2022). Then other parameters would change slightly or substantially. Therefore, it was suggested to stepwise add one parameter at a round (Whittaker & Schumacker, 2022). Besides, one or two parameters that are clearly interpretable and substantively justified were ideal for model modification, avoiding capitalization on chance (R. C. MacCallum et al., 1992), or overspecification (Fan, 2010; Fan & Hancock, 2006).

The properties of data fit to models were identified through values of χ^2/df , comparative fit index (CFI), Tucker-Lewis Index (TLI), Standardized Root Mean Square Residual (SRMR), and Root Mean Square Error of Approximation (RMSEA). χ^2/df is advised to be less than 3 (Bollen, 1989). CFI compares the model to the null model; ≥ 0.90 is acceptable, ≥ 0.95 is good. TLI is adjusted CFI, penalizing complexity; ≥ 0.90 is acceptable, ≥ 0.95 is good (Hu & Bentler, 1999). RMSEA measures error of approximation per df; ≤ 0.08 is acceptable, ≤ 0.05 is good (Brown, 2015). SRMR measures average standardized

residuals, ≤ 0.08 acceptable, ≤ 0.05 good. Cronbach's alpha is estimated to assess internal consistency (reliability) of a scale, ≥ 0.70 acceptable, ≥ 0.8 good.

Construct validity, as a unifying argument, is to evaluate the extent to which a measure assesses what it is supposed to measure (Strauss & Smith, 2009). The present research has mainly focused on convergent validity and discriminant validity. Convergent validity means correlation among different measures of the same construct (Campbell & Fiske, 1959; Fornell & Larcker, 1981; Strauss & Smith, 2009). It is reported by calculations of means of standard factor loading, composite reliability (CR), and average variance extracted (AVE) (Hair, 2010). CR (Raykov, 1997) is also known as McDonald's omega (McDonald, 2013) or congeneric reliability (Graham, 2006), and has been defined in a formula as:

$$CR = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum (1 - \lambda_i^2)}$$

Where λ represents the factor loading matrix. A construct's ability to capture variance in comparison to the amount of variance caused by measurement error is measured by its AVE (Dos Santos & Cirillo, 2023), which is calculated by the following formula:

$$AVE = (\sum \lambda^2) / n$$

CR estimates reliability based on factor loadings; ≥ 0.70 is acceptable, ≥ 0.8 is good. AVE measures convergent validity; ≥ 0.50 means good, as the latent factor explains more than 50% of item variance.

Discriminant validity (Divergent validity) indicates the extent to which a given construct differs from other constructs (Hubley, 2014). A recent definition of discriminant validity is presented by Rönkkö and Cho (2022): if the absolute value of the correlation between two measures, after adjustment for measurement error, is low enough for the measurements to be seen as measuring different constructs, then the two measures are meant to measure distinct constructs. The Fornell-Larcker criterion (Fornell & Larcker, 1981) is a method used to assess discriminant validity, which states that the square root of AVE must be more than the latent variables' correlations, $\sqrt{AVE} > \text{inter-construct correlations}$. Another recently introduced index is the heterotrait-monotrait (HTMT) ratio of correlations method that compares the correlations between constructs (heterotrait correlations) relative to the correlations within constructs (monotrait correlations) (Ab Hamid et al., 2017; Henseler et al., 2015). HTMT value ranges from 0 to 1, and < 0.90 is good in the education field (Ackermans et al., 2024; Hayat et al., 2023).

Results

Participants were recruited using an inclusive, cohort-based approach in which all students enrolled in the tennis PE module in two consecutive semesters were invited to participate. Because enrolment in this tennis module was determined by students' autonomous registration (within the university PE offering), the gender distribution of respondents largely reflects the natural composition of those enrolled cohorts rather than researcher-driven selection.

Sample 1 participants, aged 19.6 ± 3.3 , in their second year of college, were from Humanities and Social Sciences faculties. The majority of them were female ($n=300$, 84.5%). Data collection was conducted before the end of the 2024-2025-1 semester. The duration of participants' engagement in tennis PE was formed into three clusters (Table 2).

Sample 2 in their first year had similar demographic characteristics to Sample 1, in that most of them were female ($n=202$, 82.4%), and from Humanities and Social Sciences (Table 2). Forty-eight of them had some acquaintance or engagement in tennis. One hundred eighty-six of them had engaged in a ball game before learning tennis. No outliers were detected through the IQR method.

An overarching suggestion for model modification is that researchers should have sufficient substantive and theoretical knowledge to avoid overfitting (MacCallum, 1986). The following part was going to modify the model based on three methods: (1) model modification based on the original 5-factor model; (2) merging factors into a 3-factor structure and consequent modifications; and (3) modeling based on EFA provided structure and consequent modifications. Modifications were performed stepwise, such as building residual covariance path(s), deleting cross-loading item(s) or low-loading item(s), and merging highly correlated factors into a new single factor. Fit indices, convergent validity, divergent validity, and reliability would be provided for each step of modification. For Phase 1, Phase 2, and Phase 3, CFA was applied to the total data, which were converted by the z-score method and combined into one dataset. For Phase 4, data did not need to be converted because CFA and EFA focus on underlying patterns in correlations or covariances, not the raw data.

Table 2. Demographic Characteristics of Sample 1 and Sample 2

Variables		n (%)	
		Sample 1 (n=355), age 19.6±3.3	Sample 2 (n=245), age (18.9±0.7)
Gender	Male	55 (15.5)	43 (17.6)
	Female	300 (84.5)	202 (82.4)
Faculty	Animation	65 (18.3)	15 (6.1)
	Art	88 (24.8)	24 (9.8%)
	Education	29 (8.2)	45 (18.4)
	Energy	N/A	20 (8.2)
	Management	19 (5.4)	9 (3.7)
	Media	14 (3.9)	15 (6.1)
	Medical	N/A	20 (8.2)
	Secretary	25 (7.0)	8 (3.3)
	Economics	115 (32.4)	89 (36.3)
Tennis Learning Duration (semester)	One Semester	77 (21.7)	Tennis Before Yes 48 (19.6)
	Two Semesters	164 (46.2)	No 197 (80.4)
	Three Semesters	114 (32.1)	Ball Sport Before Yes 186 (75.9)
			No 59 (24.1)

Phase 1 Model Modifications Based on the Original 5-Factor Model

At the beginning, modeling motivational behavior for the present population was conducted conservatively based on the SDT-based original 5-factor model (Figure 1). CFA was conducted for the total data within the original 5-factor model.

Original 5-Factor Model for Total Data

Five-factor model fit indices were as follows: scaled $\chi^2/\text{df}=714.548/160=4.467$, RMSEA=0.091, SRMR=0.069, CFI=0.903, TLI=0.885, in which RMSEA was greater than the recommended 0.08, and TLI was less than the recommended 0.90. CR for IM, IDR, ITR, ER, and AM were 0.910, 0.913, 0.666, 0.787, and 0.848, respectively, in which CR for ITR was less than the recommended 0.7. AVE values were 0.717, 0.723, 0.344, 0.484, and 0.584, in which AVE for IRT and ER was less than 0.5, not reaching the acceptable level. Values of Cronbach's Alpha were 0.909, 0.912, 0.702, 0.783, and 0.847, all greater than the recommended 0.7 (Nunnally & Bernstein, 1994), indicating an acceptable level. Factor loadings on their own factors were mostly greater than 0.5, except for ITR2 (0.47). Fornell-Larcker and HTMT were not acceptable, indicating problems of discriminant validity (Figures 1 and 2).

In a previous study, the concern of reliability on ITR had drawn researchers' attention, with values as: 0.64 from (Ntoumanis, 2005), 0.69 from the Hong Kong sample of (Lonsdale et al., 2011), 0.69 from (Standage et al., 2005), 0.67 from the Spanish sample, and 0.66 from the Chinese sample of (Yang et al., 2019). The reliability of ER was also not fully accepted in previous studies (Hutmacher et al., 2021; Lonsdale et al., 2011; Sánchez De Miguel et al., 2017; Yang et al., 2019).

Modified Models Based on 5-Factor Structure for Total Data

First Run: Add Path of IM1~IDR1. In order to improve model fit and discriminant validity in the post-hoc session, stepwise modifications of the residual covariance path were built in this section. Residual covariance provides key insight into model fit and unaccounted misspecifications. It represents the differences between the observed data covariance matrix and the model theoretical covariance matrix. A positive value of MI means the model underestimates the observed relationships between variables. Currently, modification indices indicated MI=134.887, EPC=0.173 of IM1~IDR1. This revealed that the correlation between IM1 and IDR1 was not accounted for in the 5-factor model. Thus, a residual covariance path between them was built at the first run. Consequently, a 5-factor 1 path model was constructed, with model fit as: scaled $\chi^2/\text{df}=621.734/159=3.910$, RMSEA = 0.084 (unacceptable), SRMR = 0.069, CFI = 0.919, TLI = 0.903. CR for IM, IDR, ITR, ER, and AM were 0.910, 0.914, 0.668, 0.787, and 0.848, respectively, in which CR for ITR was less than the recommended 0.7. AVE values were 0.717, 0.726, 0.345, 0.484, and 0.584, in which AVE for IRT and ER was less than 0.5, not reaching the

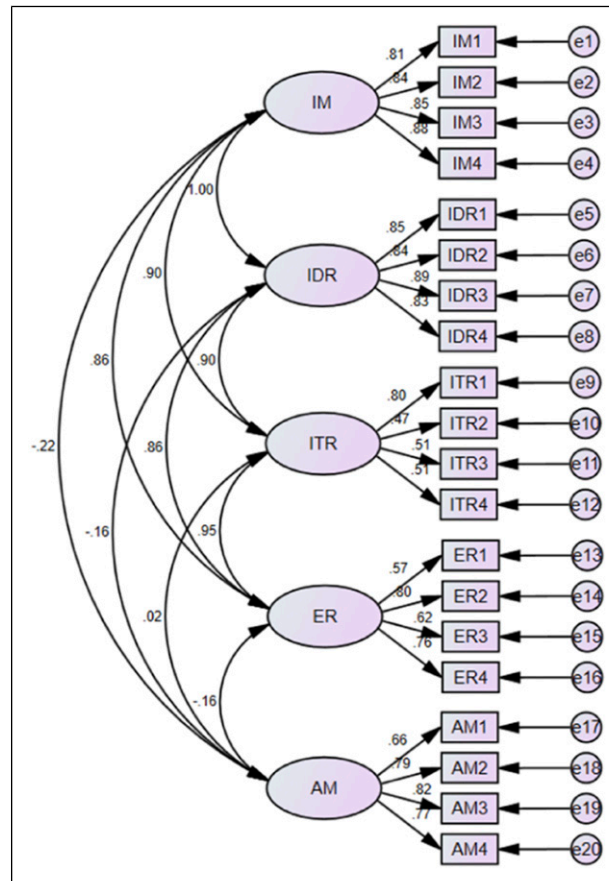


Figure 1. 5-Factor model for total data

acceptable level. Values of Cronbach's Alpha were 0.909, 0.912, 0.702, 0.783, and 0.847, all greater than 0.7, reaching an acceptable level. Fornell-Larcker and HTMT revealed that discriminant validity was not acceptable.

Previous studies validating PLOC in the PE context reported similar results about the correlations between IM and IDR, with values like 0.74 (Hutmacher et al., 2021), 0.75 (Jankauskiene et al., 2022), and 0.84 (Vlachopoulos et al., 2011).

Second Run: Add Path of ITR2~ITR4. The largest value of MI was 111.073 for ITR2 and ITR4 with EPC (0.340), representing unaccounted magnitude between observed data and the 5-factor one-path model. Therefore, a second path between ITR2 and ITR4 was built, forming a 5-factor 2-path model (Figure 2), with model fit as: scaled $\chi^2/df = 534.177/158 = 3.381$, RMSEA = 0.076, SRMR = 0.065, CFI = 0.933, TLI = 0.919, indicating an acceptable level. Cronbach's alpha was calculated based on standardized data with R (psych), still as: 0.91, 0.91, 0.70, 0.78, and 0.85.

Whereas CR, also known as construct reliability, provides more meaning of internal reliability on the modified model. Now, CR for IM, IDR, ITR, ER, and AM were 0.911, 0.917, 0.569, 0.783, and 0.847, respectively, of which the CR of ITR was still less than 0.7, not reliable. Though Cronbach's alpha of ITR (0.70) for current data was acceptable, the CR of ITR (0.569) was below the minimum acceptable threshold (0.70). Meanwhile, factor loadings of ITR2, ITR3, and ITR4 within the 5-factor 2-path model were lower than the recommended 0.5.

CR involves considerations of factor loading tune consistency. For current data, factor loadings of ITR2, ITR3, and ITR4 were different from ITR1 (0.79). 'Because I want the teacher to think that I am a good student.' This means ITR2, ITR3, and ITR4 did not have the same strength as ITR1 to represent the factor introjected regulation. Or the ITR2, ITR3, and ITR4 represented latent traits are different from ITR1, which represented orthodox, possibly revealing multiple facets (multidimensionality) for introjected regulation (Bieg et al., 2024).

In the PE motivation scope, several previous studies revealed concerns reliability and validity on ITR. ITR obtained only 9.2% of variance from observed data based on the 5-factor SDT-based motivation continuum (Hu et al., 2022). A study used a Spanish version of the PLOC (Goudas et al., 1994) through 18-20 year adolescents and demonstrated that Cronbach's alpha on ITR was 0.65, less than the recommended 0.7 (González-Cutre et al., 2014). Standage et al. (2005) reached a marginally

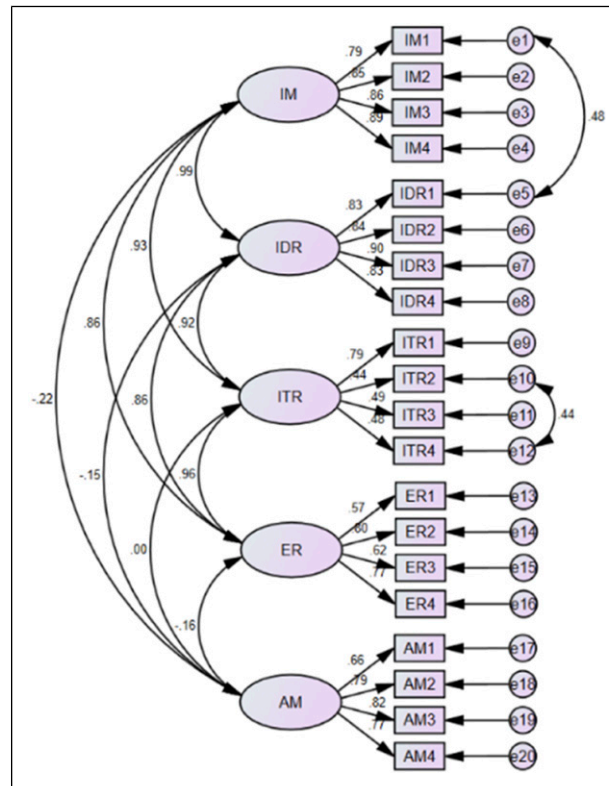


Figure 2. Five-factor two-path model

significant alpha value of 0.69 for ITR on a sample of 950 students aged from 11 to 14 years in the Midland and Southern districts of England.

A later study suggested explanations for ITR's instability. Bieg et al. (2024) argued that traditional approaches defined ITR as a single construct, ignoring the multiple facets of nature, like positive-negative and self-social approval elements, without distinction.

AVE for each construct was computed as 0.717, 0.726, 0.320, 0.484, and 0.584, slightly different from the AVEs in the 5-factor model. This is because parameters were re-estimated, such as changed factor loadings and error variances, after adding error correlations for IM1~IDR1, and ITR2~ITR4. AVE calculation relates to factor loading. AVE of ITR and ER was below the threshold (0.50), indicating that the two constructs did not capture enough variance from their indicators. AVE represents convergent validity and has a trade-off relationship with discriminant validity. On the other hand, the Larcker matrix indicated that the $\sqrt{\text{AVE}}$ of each construct ranged from 0.565 to 0.852, which were not all greater than the absolute value of inter-construct correlations, which ranged from 0.000 to 0.987 (Table 3). HTMT ratios of IM-IDR, and ITR-ER were not less than 0.90, indicating that they were not discriminant. In other words, IM-IDR ($r=0.99$) and ITR-ER ($r=0.96$) showed too high correlations (Table 3).

IM1 'Because physical education is fun.' might have content overlaps with responders' experience and understanding with IDR1 'Because I want to learn PE/sports skills.' IDR is defined as an almost internalized autonomous motivation (Deci & Ryan, 2008). Responders might solely identify 'physical education is fun' and 'I want to learn PE/sports skills' as reflecting the same degrees of approving/liking PE. In natural identification regulation development, one would explicitly realize the importance/value of an activity, but full self-compatibility check or full integration is not necessary at this phase (Ryan & Deci, 2017).

For measurement practice, it seems that more distinct expressions of IDR1 apart from IM1 could possibly be a valuable proposition in the future to improve discriminant validity between IM and IDR. For instance, formative items could be an option to improve distinctions between IM and IDR. Formative items cause, define, or form the construct, and they are not necessarily correlated. The contributions and concreteness from formative items would provide an alternative option for improving distinctions between constructs.

Table 3. Fornell-Larcker Matrix & HTMT Ratio of One Round-Modified 5-Factor Model

	IM	IDR	ITR	ER	AM
IM	0.847	1.003	0.714	0.837	0.197
IDR	0.987	0.852	0.748	0.835	0.120
ITR	0.929	0.925	0.565	0.936	0.232
ER	0.863	0.856	0.963	0.696	0.119
AM	-0.220	-0.152	0.000	-0.159	0.764

Note. Values within diagonal are the square root of AVE, the lower triangle values are construct correlations, and the upper triangle values are HTMT ratios.

Phase 2 Modeling SDT-Based 4-Factor Model and Necessary Modifications for Total Data

Making models toward parsimonious ones is considered to be a consensual methodology in the scale evaluation field. Though there was a simplified conceptual framework in the history of SDT development (Deci & Ryan, 2008). At this phase, the modification pattern would not directly go to 3-factor structures. Before discussing 3-factor structures, a serial four-factor structure based on present observed data would be introduced for exploration.

Original. The SDT-based 4-factor model was formulated through merging IM and IDR into one single factor, F1. Its fit indices were as follows: scaled $\chi^2/df = 719.168/164 = 4.385$, RMSEA = 0.091, SRMR = 0.070, CFI = 0.902, TLI = 0.886. Cronbach's alpha for each construct was 0.954, 0.702, 0.783, and 0.847, all reaching acceptable reliability. CR for F1, ITR, ER, and AM were 0.954, 0.666, 0.787, and 0.848, all reaching good factor strength and tune. AVE for each construct was 0.721, 0.344 (lower than 0.50), 0.484, and 0.584, in which ITR did not obtain acceptable convergent validity, while ER was at a marginal construct convergent level.

First run. Because AVE did not all reach an acceptable level, a meaningful operation was MI calculation, obtaining values 119.799 with an EPC value of 0.156 for IM1~IDR. Accordingly, a path was built and the model was formulated as a new 20-item 1-path 4-factor model. The current model fit was as follows: scaled $\chi^2/df = 632.525/163 = 3.881$, RMSEA = 0.084, SRMR = 0.069, CFI = 0.917, TLI = 0.903. Cronbach's alpha for each construct was 0.954, 0.702, 0.783, and 0.847, all reaching acceptable reliability. CR for F1, ITR, ER, and AM were 0.947, 0.667, 0.787, and 0.848, all reaching good factor strength and tune. AVE for each construct was 0.716, 0.345 (lower than 0.50), 0.484, and 0.584, in which ITR did not obtain acceptable convergent validity, while ER was at a marginal construct convergent level.

Second run. According to the last two sets of model fit indices, and especially of the convergent validity from AVE indications. The items made AVE at a low level were possibly within the ITR or ER construct. Therefore, item deletion modifications are performed at this run. The deleted item was determined from the low-loading item of the last run. The lowest value within the standardized factor loading list was ITR2. A 19-item 1-path 4-factor model was developed, with model fit as: scaled $\chi^2/df = 485.552/145 = 3.881$, RMSEA = 0.077, SRMR = 0.061, CFI = 0.935, TLI = 0.923, all at an acceptable level. Cronbach's alpha for each construct was 0.954, 0.612, 0.783, and 0.847, in which the value of ITR decreased from 0.702 to 0.612. CR for F1, ITR, ER, and AM were 0.947, 0.616, 0.786, and 0.848, in which the value of ITR decreased from 0.667 to 0.616. Similarly to the last run, AVE for each construct was 0.717, 0.362 (lower than 0.50), 0.484, and 0.584, in which ITR did not obtain acceptable convergent validity, while ER was still at a marginal construct convergent level. Conclusively, the strategy of deleting ITR2 did not seem fully appropriate for increasing model goodness both in reliability and validity.

Phase 3 Developing SDT-Based 3-Factor Model and Necessary Modifications for Total Data

Deci and Ryan (2008) defined that autonomous motivation comprised of intrinsic motivation and identified regulation, while controlled motivation included introjected regulation and external regulation. Highly correlated variables could be merged into a single variable (Kyriazos & Poga, 2023). For the present research, the two pairs of highly correlated factors, IM-IDR and ITR-ER, were combined into new factors, named Autonomous Motivation and Controlled Motivation. Thus, the 3-Factor model comprised three structures as Autonomous Motivation (ATM), Controlled Motivation (CTM), and Amotivation (Figure 3).

The model fit of the SDT-based 3-factor model was as follows: scaled $\chi^2/df = 741.694/167 = 4.441$, RMSEA = 0.092, SRMR = 0.073, CFI = 0.898, TLI = 0.884. Cronbach's alpha for each construct was 0.954, 0.847, and 0.847, all reaching good reliability. CR for ATM, CTM, and AM were 0.954, 0.837, and 0.848, all reaching good factor strength and tune. AVE for each construct was 0.721, 0.401 (lower than 0.50), and 0.584, in which CTM did not obtain acceptable convergent validity.

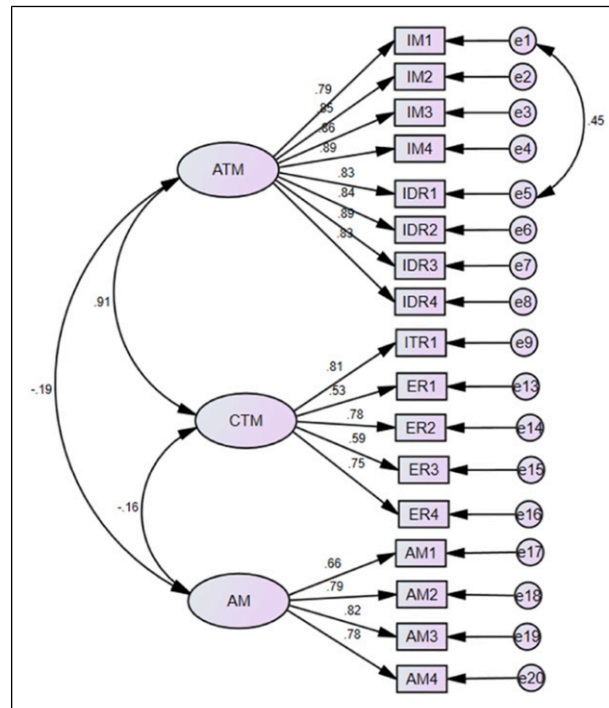


Figure 3. Modified SDT based 3-factor model

First Run: Add Path of IM1~IDR1. The largest MI value was 120.148 for IM1~IDR1, with EPC (0.157), which was the same error correlation as the former 5-factor model ever emerged. Based on the SDT theory foundation and former empirical evidence, a residual covariance path between IM1~IDR1 was built. Then an SDT-based 3-factor 1-path model was developed with model fit as: scaled $\chi^2/df=655.809/166=3.951$, marginal acceptable value of RMSEA=0.085, SRMR=0.073, CFI=0.912, TLI=0.900. Cronbach's alpha for each construct was 0.954, 0.847, and 0.847. CR for ATM, CTM, and AM were 0.947, 0.801, and 0.848. AVE for each construct was 0.716, 0.402 (still lower than 0.50), and 0.584. AVE value for CTM indicated that these indicators within CTM were still not converging at the minimum requirement. Because AVE evaluates the item loading on its construct. Therefore, the model modification strategy shifted attention to removing low-loading item(s).

Second Run: Removing ITR2. The lowest factor loading was 0.467 for ITR2. After removing ITR2, a 19-item 1-path 3-factor model was developed, with model fit: scaled $\chi^2/df = 420.089/131=3.207$, RMSEA=0.076, SRMR=0.055, CFI=0.941, TLI=0.931. CR for constructs were 0.948, 0.803, and 0.847. AVE for constructs were 0.717, 0.421 (still less than 0.5), 0.584. Cronbach's alpha for each construct was 0.954, 0.835, and 0.847.

Third Run: Removing ITR3. The lowest factor loading within the 19-item 1-path 3-factor model was 0.474 for ITR3. After removing ITR3, an 18-item 1-path 3-factor model was developed, with model fit: scaled $\chi^2/df=510.038/148=3.446$, RMSEA=0.079, SRMR=0.063, CFI=0.931, TLI=0.920. CR for constructs were 0.948, 0.797, and 0.847. AVE for constructs were 0.717, 0.451 (not greater than 0.5), and 0.584. Cronbach's alpha for each construct was 0.954, 0.834, and 0.847.

Fourth Run: Removing ITR4. The lowest factor loading within the 18-item 1-path 3-factor model was 0.472 for ITR4. After removing ITR4, a 17-item 1-path 3-factor model was developed, with model fit: scaled $\chi^2/df = 322.102/115=3.446$, RMSEA=0.072, SRMR=0.047, CFI=0.952, TLI=0.944. CR for constructs were 0.948, 0.807, 0.847. AVE for constructs were as 0.717, 0.491, 0.584. Cronbach's alpha for each construct was 0.954, 0.828, and 0.847. Factor loadings were all greater than 0.5, indicating the remaining items had enough strength to represent the three factors.

Though AVE for CTM was marginally met, discriminant validity was examined using the Fornell-Larcker matrix and the HTMT ratio. On the one hand, values of the HTMT ratio were all less than 0.90 (Table 4), indicating acceptable discriminant validity among the three constructs. On the other hand, for Fornell-Larcker matrix, $\sqrt{AVE}$ of each construct ranged from 0.700 to 0.847 were not all greater than the absolute value of inter-construct correlations ranged from 0.165 to 0.914 (Table 4). Implications derived from the HTMT ratio seemed to be contradictory to the Fornell-Larcker matrix.

In fact, previous studies clarified that HTMT demonstrates superior performance compared to the Fornell-Larcker criterion (Franke & Sarstedt, 2019; Henseler et al., 2015; Radomir & Moisescu, 2019). For example, Franke and Sarstedt

Table 4. Fornell-Larcker Matrix and HTMT Ratio of SDT-Based 3-Factor Model

	ATM	CTM	AM
ATM	0.847	0.875 [0.830, 0.915]	0.153 [0.073, 0.253]
CTM	0.914	0.700	0.118 [0.077, 0.189]
AM	-0.187	-0.165	0.764

Note. ATM=Autonomous Motivation, CTM=Controlled Motivation, AM=Amotivation; Values within diagonal are the square root of AVE, the lower triangle values were constructs correlation, and the upper triangle values are HTMT ratios; the values in square brackets are the 95% bias-corrected and accelerated confidence intervals of the HTMT statistic.

(2019) commented that Fornell-Larcker was a heuristic rather than a statistical test, and focused on item loadings of AVE but ignored CR. Besides, (Rönkkö & Cho, 2022) evaluated Fornell-Larcker as a conservative test.

For the present sample's specification, the Fornell-Larcker Matrix did not support discriminant validity. Though the values of HTMT ratios were all less than 0.9, the value between ATM and CTM was greater than 0.85 (a more stringent threshold). In addition, confidence intervals that were derived from 10000 bootstrap samples were introduced to warrant an error probability (Hair et al., 2021). The upper bound of confidence intervals of HTMT (0.915) for ATM-CTM was greater than 0.9, indicating ATM and CTM were almost distinct at marginal criteria. HTMT ratio and bootstrapped HTMT were all less than 0.85, revealing AM was distinct from ATM and CTM. Generally, constructs within the 17-item 1-path SDT-based 3-factor model obtained marginal discriminant validity, trading off 3 ITR items' information loss.

To sum up, the first problem about IM1~DIR1 high correlation was fixed by merging IM and IDR into a new construct, ATM. Then, the AVE for ATM is acceptable. But the second problem, the low factor loading of ITR2, ITR3, and ITR4, caused unacceptable AVEs for CTM, which was still not fully fixed by removing ITR2, ITR3, and ITR4; it only achieved a marginal value and caused information loss. A controversy and a crucial joint emerged from the convergent validity (AVE) of ITR within the 5-factor model and of CTM within the SDT-based 3-factor model. Furthermore, it seems that a more exploratory way is necessary to reexamine the factorial structure of observed data, especially focusing on ITR items.

Phase 4 Model Development Based on EFA Provided Structures

EFA on Sample 1

Though the two samples were representative of the population, they were independent and natural. In this phase, EFA and CFA tailored the analysis to the two samples. The factorial structure and number of factors were derived from EFA on Sample 1. A series of indices were calculated in R (psych, lavaan, and semTools), including (1) indices derived from EFA, and (2) CR, AVE, model fit, Fornell-Larcker Matrix, and HTMT ratio, and modification indices of CFA.

For Sample 1, no missing data were identified through SPSS, and 1 case was identified as an outlier (case 245), eliminated before EFA. Next, Bartlett's test of sphericity was significant as the approximate Chi-Square was 5779.716, $p < 0.001$, and KMO was 0.940, confirming the data were suitable for factor analysis. Then, eigenvalues before rotation were computed in the process of EFA, as 9.424, 2.191, and 0.539. Parallel analysis and scree plot (Figure 4) revealed that the optimal factor number for explained variance in observed data was three, which was just larger than the simulated data.

EFA on Sample 1 firstly revealed that the optimal number of factors for the present samples was three (Figure 4, Table 5). Pattern matrix of Sample 1 detected two cross-loading items, ITR3 and ER4. But deleting items is more radical than building a residual covariance path and losing item information. Thus, residual covariance path building was primarily applied. Before model modification, cross-loading problems for ITR3 and ER4 were discussed. Apart from former studies concerned with ITR reliability (González-Cutre et al., 2023; Goudas et al., 1994; Hu et al., 2022; Standage et al., 2005), and ignoring multiple facets of ITR (Bieg et al., 2024), there was also another study (Haerens et al., 2015) that reported the issue of cross-loading (0.40) and low loading (0.42).

Because we use a two-step approach of EFA on Sample 1 and CFA on Sample 2, the pattern matrix detected problematic low-loading items and cross-loading items will still be kept before CFA on another sample. Thus, path building was the primary strategy used for modifying the model and checking the indices improvement at the CFA stage. Till met with construct validity problems, like unacceptable indices from AVE (convergent validity) and Fornell-Larcker accompanied with HTMT (discriminant validity), item deletion within construct detection level, and reorganizing items across constructs were brought into consideration (ever used in Phase 3 for merging factors, and would be used in the current Phase based on

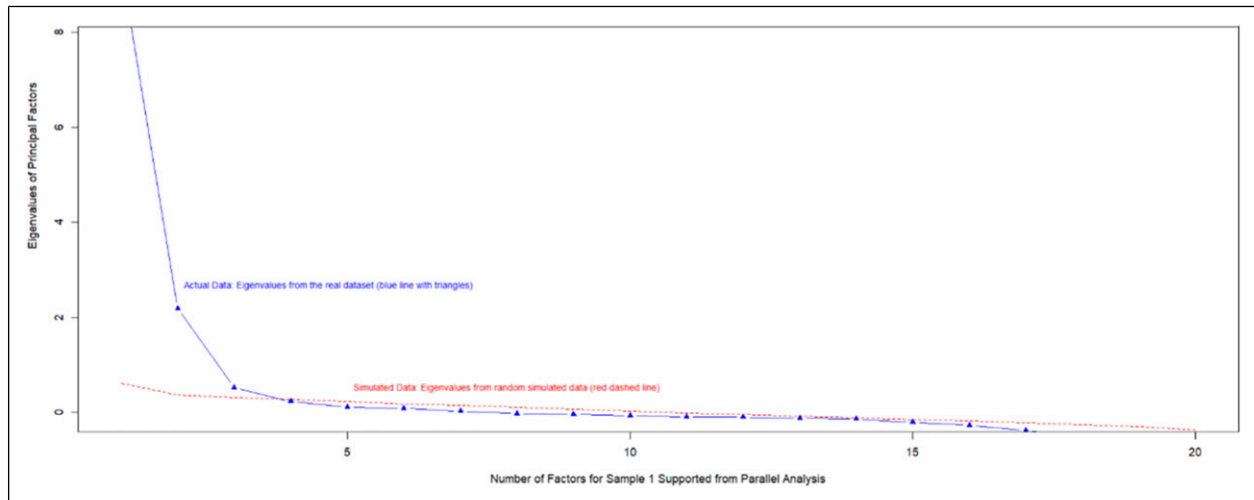


Figure 4. Parallel analysis scree plot of sample 1

item correlation matrix). Cautions on deleting item information were based on information loss caused by changes to other elements of a model.

CFA on Sample 2

Original. EFA on sample 1 derived structure as (in syntax): $APM \sim IM1 + IM2 + IM3 + IM4 + IDR1 + IDR2 + IDR3 + IDR4 + ITR1 + ER2 + ER4$, $AVM \sim ER1 + ITR2 + ER3 + ITR4$, $AM \sim AM1 + AM2 + AM3 + AM4 + ITR3$. This structure was applied with CFA on Sample 2 data, achieving model fit as: scaled $\chi^2/df = 378.146/167 = 2.264$, RMSEA = 0.086, SRMR = 0.104, CFI = 0.894, TLI = 0.879. CR for constructs were 0.937, 0.716, and 0.724, all greater than 0.70, considered acceptable. AVE for constructs were 0.581, 0.404, and 0.370, of which the AVEs for AVM and AM were less than 0.5,

Table 5. Pattern Matrix of Sample 1
(PAF With Oblimin Rotation)

Variables	Sample 1		
	F1	F2	F3
IM1	0.825		
IDR1	0.873		
ITR1	0.719		
ER1		0.782	
AM1			0.701
IM2	0.806		
IDR2	0.774		
ITR2		0.708	
ER2	0.625		
AM2			0.815
IM3	0.982		
IDR3	0.983		
ITR3	0.354		0.365
ER3		0.530	
AM3			0.736
IM4	0.866		
IDR4	0.864		
ITR4		0.681	
ER4	0.486	0.333	
AM4			0.643

Note. The values in bold demonstrate cross-loadings.

considered unacceptable. Cronbach's alpha for each construct was 0.935, 0.690, and 0.713, of which Cronbach's alpha for AVM was not greater than 0.7, which was considered marginally acceptable. AVE for AVM and AM had problems (less than the recommended 0.5), then the Fornell-Larcker matrix and the HTMT ratio would not compute. Next, detections of items loading on AVM or AM are performed for deletion.

First run. Thus, ITR3 was detected with a low loading value of 0.135. After deleting ITR3, a 19-item 3-factor model was developed, with model fit as: scaled $\chi^2/df = 304.366/149 = 2.043$, RMSEA=0.081, SRMR=0.071, CFI=0.914, TLI=0.902, in which RMSEA was not less than 0.08, but considered marginally acceptable. The other model fit indices were less than the threshold and acceptable. CR for constructs were 0.936, 0.726, and 0.822, in which the CR of F3 increased from 0.709 to 0.822 by deleting ITR3. Cronbach's alpha for each construct was 0.935, 0.670, and 0.841. AVE for constructs were 0.581, 0.404, and 0.558, in which AVE for AM increased from 0.370 to 0.558 by deleting ITR3. Up to now, AVE for AVM was not acceptable. The next round of detection focused on the loading of items of the AVM within the 19-item 3-factor model. ER1 loading on AVM was 0.477 currently. Then, for considering another side, MI for IM1~IDR1 was 51.307 with EPC 32.325.

Second run. Path building for IM1~IDR1 within F1 would not improve AVE for AVM. Therefore, deleting ER1 applied and developed an 18-item 3-factor model (Figure 5). Now the model fit indices were: scaled $\chi^2/df = 254.218/132 = 1.926$, RMSEA=0.079, SRMR=0.066, CFI=0.925, TLI=0.913, all acceptable. CR for constructs were 0.936, 0.717, and 0.822, all greater than 0.7, considered acceptable. Cronbach's alpha for each construct was 0.935, 0.678, and 0.841, in which Cronbach's alpha for AVM slightly increased from 0.670 to 0.678 compared with the last run. Now, AVEs for constructs were 0.581, 0.460, and 0.558, and the AVE for AVM was marginally acceptable.

The Fornell-Larcker matrix and the HTMT ratio were computed. The Fornell-Larcker matrix indicated that $\sqrt{\text{AVE}}$ of each construct ranged from 0.678 to 0.762, all of which were greater than the absolute value of inter-construct correlations, which ranged from 0.089 to 0.380 (Table 6). HTMT ratios were all less than 0.85, indicating that they were discriminant (Table 6).

Summing up Phase 4, parallel analysis and pattern matrix of EFA on sample 1 obtained an optimal number of factors, which were three. The pattern matrix revealed factorial structure, and it was introduced to Sample 2 for CFA, obtaining acceptable levels of reliability, model fit, convergent validity, and divergent validity. Previous studies raised concerns about the reliability and validity of ITR items. Actions of ITR are controlled by rules set by oneself; in other words, the aim of ITR

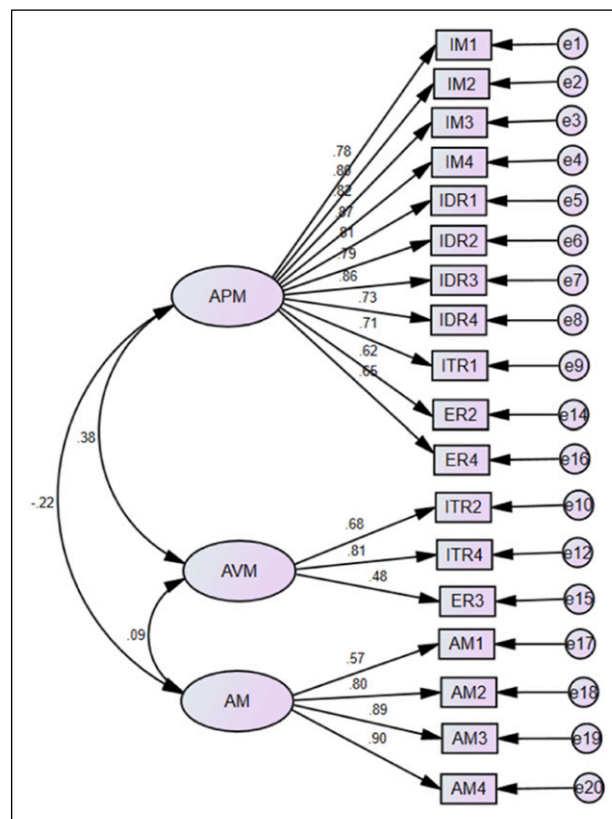


Figure 5. 3-Factor 18-item model from EFA of sample 1 and CFA on sample 2

Table 6. Fornell-Larcker Matrix and HTMT Ratio of AGT-SDT-Based 3-Factor 19-Item 1-Path Model

	APM	AVM	AM
APM	0.762	0.470	0.206
AVM	0.380	0.678	0.063
AM	-0.223	0.089	0.747

Note. APM=Approach Motivation, AVM=Avoidance Motivation, AM=Amotivation; Values within diagonal are the square root of AVE, the lower triangle values are construct correlations, and the upper triangle values are HTMT ratios; the values in square brackets are the 95% bias-corrected and accelerated confidence intervals of the HTMT statistic.

action is the avoidance of negative self-evaluation (Goudas et al., 1994). The latent construct ITR is named in objective (third-person) perspectives of external value internalization. While approach-avoidance is named in the first person to describe the direction of motivational behaviors. Approach motivation is about an individual's behavior being directed by a positive or desirable event or possibility, while avoidance motivation is about behaviors directed by a negative or undesirable event or possibility (Elliot, 1999). The items of intrinsic motivation (item 1, item 6, item 11, item 16), identified motivation (item 2, item 7, item 12, item 17), introjected regulation (item 3), and external regulation (item 9, item 19) describe positive directions (or stimuli, or events, or possibilities) of one's behavior (Figure 5).

In fact, the Approach and Avoidance constructs are conversed and deducted from constructs of IM, IDR, ITR, and ER by ignoring the four degrees of internalization and emphasizing the dichotomous orientation of motivation. Approach-avoidance is a rudiment of AGT, not emphatically reported in the motivation field. But it is a basic rationale to understand motivational behavior. Amotivation was still consistent with the original PLOC. Therefore, for the current observed data, we renamed the EFA-suggested 3 factors as Approach Motivation, Avoidance Motivation, and Amotivation, to give implications of options in real-world research practice.

Conclusions to the three Phases analysis. Generally, Phase 1 used post-hoc model modification, while Phase 3 and Phase 4 used the dimension reduction method appropriately when construct validity was not achieved in the 5-factor models.

Discussion

The present study was designed to evaluate the psychometric structure of a Chinese-language PLOC instrument in a university tennis PE context. Accordingly, our central contribution is measurement-focused: EFA of Phase 4 explored the latent structure, and CFA of Phase 4 tested theoretically plausible alternatives. This combined approach allowed us to examine whether the SDT-based regulations represented by PLOC items functioned as distinct factors in this setting, and whether a more parsimonious structure offered improved psychometric performance.

Within CFA, a residual covariance between IM1 and IDR1 was introduced to account for local item dependence. This adjustment is most plausibly explained by overlap in item wording and respondent interpretation, rather than by substantive theoretical coupling that would warrant a new construct. In other words, the residual link was treated as an item-level refinement to improve model representation, consistent with the broader pattern that intrinsic motivation and identified regulation were empirically difficult to separate in this sample.

Across competing models, the three-factor structure demonstrated superior fit compared with both the original five-factor model and the four-factor alternative. This finding aligns with arguments that PLOC regulations may collapse into higher-order groupings in some PE contexts, particularly when intrinsic and identified endorsements are experienced as closely intertwined in practical participation settings. A similar three-factor structure was proposed by Hutmacher et al. (2021), comprising autonomous motivation (intrinsic plus identified), controlled motivation (introjected plus external), and amotivation. However, our results diverge from Hutmacher et al. (2021) in that the three-factor solution outperformed the

five-factor model in the current dataset. Rather than interpreting this discrepancy as a contradiction of SDT theory, it is more appropriately understood as evidence of context-sensitive measurement behavior. Differences in activity domain (tennis PE), language adaptation, and cohort characteristics may alter how students distinguish between “enjoyment” and “valuing” statements, even when the underlying theoretical continuum remains intact.

Evidence from the five-factor solution supports this interpretation. Intrinsic motivation and identified regulation exhibited an extremely high association ($r = .98$), consistent with prior studies reporting very strong correlations between these regulations in PE-related samples (Ferriz et al., 2015; Hutmacher et al., 2021; Lonsdale et al., 2011; Murcia et al., 2009; Sánchez De Miguel et al., 2017; Wang et al., 2009; Wolf et al., 2022; Yang et al., 2019). Convergent and discriminant validity checks reinforced the same conclusion: both the Fornell-Larcker criterion and HTMT ratios indicated inadequate discriminant validity between intrinsic motivation and identified regulation in the four-factor and five-factor specifications. By contrast, amotivation remained clearly distinguishable from the other regulations. Taken together, these results support a parsimonious reorganization in which intrinsic and identified items form a single autonomous motivation factor, introjected and external items form controlled motivation, and amotivation remains separate.

Although SDT is the guiding framework for the present measurement work, PE is also an achievement setting where AGT constructs are often examined alongside SDT variables in the broader literature. We reference SDT–AGT integration only to situate PLOC measurement within this wider body of PE motivation scholarship (Duda, 2013; Standage et al., 2003). However, the present study did not measure achievement goals, motivational climate, or other AGT indicators, and therefore cannot test integrated SDT–AGT process claims. Future studies should treat SDT–AGT integration as an a priori design feature by collecting validated PLOC indicators together with AGT measures (goal orientations and motivational climate) and then testing measurement invariance and theoretically specified relations within a single structural model.

Limitation

Although we followed recommended translation and adaptation procedures and documented item-level revisions, the present study did not directly test cross-language measurement invariance because the English and Chinese versions were not administered in parallel within the same sampling frame. Accordingly, the adapted items should be interpreted as a Chinese-language operationalization of the PLOC constructs tailored to a tennis PE context, rather than as evidence of strict comparability with all prior English-language implementations.

Conclusion

A 17-item 1-path 3-factor model obtained acceptable indices and converged with a simplified SDT continuum of Autonomous Motivation, Controlled Motivation, and Amotivation. The EFA-CFA analysis procedure derived an 18-item 3-factor model, performed with acceptable quality and developmental adaptations in the present population. Generally, the present population was appropriate to the parsimonious SDT-based three-factor model, as well as the SDT-AGT-based three-factor model.

Suggestions for Future Research

For future research, firstly, at the item revision level, more work can be done to optimize the 3-factor model-based PLOC, as the previous studies concerned high correlations between IM and IDR. Secondly, for scale selection, future studies can focus on concurrent validity between SDT-framed instruments like PLOC (Goudas et al., 1994), and Behavioral Regulation in Exercise Questionnaire (BREQ) (Mullan et al., 1997), and AGT-framed instruments like AGQ (Elliot & McGregor, 2001) to help researchers detect which instruments are more suitable for their participants.

Then it is still necessary to test the fit, validity, and reliability of the 5-factor model-based PLOC in a larger sample size, especially in a cross-cultural context. Specifically, linguistic equivalence is an important topic when a scale is translated, as it is combined with cultural differences. In practice, an expert panel that translates the scale and evaluates the translated scale might waver between the literal and free translation methods, impairing equivalence with the original scale. Thus, experimental studies could examine the power of different versions of scales, like bilingual and monolingual scales, for Chinese university students. The main issue that the present research is greatly concerned with is the discriminant validity among pairs of IM-IDR and ITR-ER. The five degrees of motivation within PLOC had challenged the researcher who aimed to validate PLOC in his population. Researchers of the present study reminded that scale literal structure and emotional expressions might benefit scale discriminant validity. For instance, experimental studies could design a testlet-based scale and a traditional scale to determine the improvement in literal-structured scale discriminant validity, in which a testlet means adding an explanatory paragraph before each block to ease responders' understanding and differentiate constructs.

Lastly, several adjacent scale testing areas would promote PLOC development such as nomological validity, predictive validity, and measurement invariance across gender, faculty, sport project, etc.

Acknowledgement

Liu would like to acknowledge the great support of the management and administration at Weifang Institute of Technology.

ORCID iDs

Changjie Liu  <https://orcid.org/0009-0000-1087-9491>

Nung Kion Lee  <https://orcid.org/0000-0001-9651-5225>

Kee-Man Chuah  <https://orcid.org/0000-0001-7499-5889>

Ethical Considerations

The studies involving human participants were reviewed and approved by the Ethics Committee of Universiti Malaysia Sarawak. Before conducting testing and distributing survey questionnaires, each participant would receive an informed consent form. If the participants confirmed that they agreed to participate in the survey, they would continue to fill in the questionnaire; If not, the participant would immediately stop responding.

Consent to Participate

We confirmed that all research was performed in accordance with the Declaration of Helsinki, and we obtained informed consent from all participants.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors would like to acknowledge Universiti Malaysia Sarawak (UNIMAS) for the support provided towards this publication.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Ab Hamid, M. R., Sami, W., & Mohmad Sidek, M. H. (2017). Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion. *Journal of Physics: Conference Series*, 890(1), 012163. <https://doi.org/10.1088/1742-6596/890/1/012163>
- Abós, Á., Sevil, J., Martín-Albo, J., Aibar, A., & García-González, L. (2018). Validation Evidence of the Motivation for Teaching Scale in Secondary Education. *The Spanish Journal of Psychology*, 21(e9), 1–12. <https://doi.org/10.1017/sjp.2018.11>
- Ackermans, K., Bakker, M., Gorissen, P., Van Loon, A., Kral, M., & Camp, G. (2024). Development and validation of a test for measuring primary school students' effective use of of ICT: The ECC-ICT test. *Journal of Computer Assisted Learning*, 40(3), 960–972. <https://doi.org/10.1111/jcal.12924>
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84(3), 261–271. <https://doi.org/10.1037/0022-0663.84.3.261>
- Auerswald, M., & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. *Psychological Methods*, 24(4), 468–491. <https://doi.org/10.1037/met0000200>
- Bartlett, M. S. (1950). TESTS OF SIGNIFICANCE IN FACTOR ANALYSIS. *British Journal of Statistical Psychology*, 3(2), 77–85. <https://doi.org/10.1111/j.2044-8317.1950.tb00285.x>
- Bieg, S., Thomas, A. E., Spreitzer, C., & Müller, F. H. (2024). Exploring the Complexity of Introjected Regulation in Self-Determination Theory. *International Journal of Teaching and Learning Sciences*, 1–10. <https://www.cmjpublishers.com/volume-1-issue-2-ijtls/>
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons. <https://doi.org/10.1002/9781118619179>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second edition). The Guilford Press. <https://psycnet.apa.org/record/2015-10560-000>
- Byrne, B. M. (2016). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming, Third Edition* (0 edn). Routledge. <https://doi.org/10.4324/9781315757421>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105. <https://doi.org/10.1037/h0046016>
- Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

- Chen, S., & Solmon, M. (2024). Fostering Motivation in Physical Education to Promote Learning and Achievement. *Kinesiology Review*, 13(2), 207–217. <https://doi.org/10.1123/kr.2024-0001>
- Costello, A. B., & Osborne, J. W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis. *Practical Assessment Research & Evaluation*, 10, 1–9. https://legacyfileshare.elsevier.com/promis_misc/teln-best-practices-in-exploratory-factor-analysis.pdf
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer. <https://link.springer.com/book/10.1007/978-1-4899-2271-7>.
- Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology / Psychologie Canadienne*, 49(3), 182–185. <https://doi.org/10.1037/a0012801>
- DeVellis, R. F., & Thorpe, C. T. (2022). *Scale development: Theory and applications* (Fifth edition). Sage. <https://www.scribd.com/document/966516914/93377>
- Dos Santos, P. M., & Cirillo, M. Â. (2023). Construction of the average variance extracted index for construct validation in structural equation models with adaptive regressions. *Communications in Statistics - Simulation and Computation*, 52(4), 1639–1650. <https://doi.org/10.1080/03610918.2021.1888122>
- Duda, J. L. (2013). The conceptual and empirical foundations of Empowering Coaching™: Setting the stage for the PAPA project. *International Journal of Sport and Exercise Psychology*, 11(4), 311–318. <https://doi.org/10.1080/1612197X.2013.839414>
- Duda, J. L., & Appleton, P. R. (2016). Empowering and Disempowering Coaching Climates: Conceptualization, Measurement Considerations, and Intervention Implications. In *Sport and Exercise Psychology Research* (pp. 373–388). Elsevier. <https://doi.org/10.1016/B978-0-12-803634-1.00017-0>
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34(3), 169–189. https://doi.org/10.1207/s15326985ep3403_3
- Elliot, A. J., & McGregor, H. A. (2001). A 2 X 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501–519. <https://doi.org/10.1037/0022-3514.80.3.501>
- Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press. <https://global.oup.com/academic/product/exploratory-factor-analysis-9780199734177?cc=tw&lang=en&>.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Fan, W. (2010). Impact of Measurement Model Modification on Structural Parameter Integrity When Measurement Model is Misspecified. *Journal of Modern Applied Statistical Methods*, 9(1), 103–115. <https://doi.org/10.22237/jmasm/1272687060>
- Fan, W., & Hancock, G. R. (2006). Impact of Post Hoc Measurement Model Overspecification on Structural Parameter Integrity. *Educational and Psychological Measurement*, 66(5), 748–764. <https://doi.org/10.1177/0013164405285908>
- Ferriz, R., González-Cutre, D., & Sicilia, Á. (2015). Revisión de la Escala del Locus Percibido de Causalidad (PLOC) para la Inclusión de la Medida de la Regulación Integrada en Educación Física. *Revista de Psicología Del Deporte*, 24(2), 329–338. <https://www.redalyc.org/articulo.oa?id=235141413017>
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Franke, G., & Sarstedt, M. (2019). Heuristics versus statistics in discriminant validity testing: A comparison of four procedures. *Internet Research*, 29(3), 430–447. <https://doi.org/10.1108/IntR-12-2017-0515>
- García-González, L., García-Cazorla, J., & Abós, Á. (2025). Motivational Climates from Achievement Goal Theory: Task-Oriented Climate as a Motivational Tool. In L. García-González, K. De Cocker, & D. González-Cutre (Eds.), *Motivation in Physical Education* (pp. 245–258). Springer. https://doi.org/10.1007/978-3-031-86908-2_12
- González-Cutre, D., Brugarolas-Navarro, M., Beltrán-Carrillo, V. J., & Jiménez-Loaisa, A. (2023). The frustration of novelty and basic psychological needs as predictors of maladaptive outcomes in physical education. *Phys. Educ. Sport Pedagogy*, 30(1), 1–14. <https://doi.org/10.1080/17408989.2023.2167969>
- González-Cutre, D., Sicilia, Á., Beas-Jiménez, M., & Hagger, M. S. (2014). Broadening the trans-contextual model of motivation: A study with Spanish adolescents. *Scandinavian Journal of Medicine & Science in Sports*, 24(4), e306–e319. <https://doi.org/10.1111/sms.12142>
- Goudas, M., Biddle, S., & Fox, K. (1994). Perceived locus of causality, goal orientations, and perceived competence in school physical education classes. *British Journal of Educational Psychology*, 64(3), 453–463. <https://doi.org/10.1111/j.2044-8279.1994.tb01116.x>
- Graham, J. M. (2006). Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability: What They Are and How to Use Them. *Educational and Psychological Measurement*, 66(6), 930–944. <https://doi.org/10.1177/0013164406288165>
- Haerens, L., Aelterman, N., Vansteenkiste, M., Soenens, B., & Van Petegem, S. (2015). Do perceived autonomy-supportive and controlling teaching relate to physical education students' motivational experiences through unique pathways? Distinguishing between the bright and dark side of motivation. *Psychology of Sport and Exercise*, 16(3), 26–36. <https://doi.org/10.1016/j.psychsport.2014.08.013>

- Hair, J. F. (Ed.), (1995). *Multivariate data analysis: With readings* (4. ed.). Prentice Hall International. https://books.google.com/books/about/Multivariate_Data_Analysis_with_Readings.html?id=vtoKAAACAAJ
- Hair, J. F. (Ed.), (2010). *Multivariate data analysis* (7th ed.). Prentice Hall. <https://www.drmishikantjha.com/papersCollection/MultivariateDataAnalysis.pdf>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-80519-7>
- Hambleton, R. K., & Patsula, L. (1998). Adapting Tests for Use in Multiple Languages and Cultures. *Social Indicators Research*, 45(1–3), 153–171. <https://doi.org/10.1023/A:1006941729637>
- Hayat, A. A., Shateri, K., Kamalian Fard, S., Sabzi Shahr Babak, E., & Faraji Dehsorkhi, H. (2023). Psychometric properties of the persian version of the physician teaching self-efficacy questionnaire. *BMC Medical Education*, 23(1), 163. <https://doi.org/10.1186/s12909-023-04130-6>
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Hu, T., Zhang, M., Liu, H., Liu, J., Pan, S., Guo, J., Tian, Z., & Cui, L. (2022). The influence of “small private online course + flipped classroom” teaching on physical education students’ learning motivation from the perspective of self-determination theory. *Frontiers in Psychology*, 13, 938426. <https://doi.org/10.3389/fpsyg.2022.938426>
- Hubley, A. M. (2014). Discriminant Validity. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 1664–1667): Springer. https://doi.org/10.1007/978-94-007-0753-5_751
- Hutmacher, D., Eckelt, M., Bund, A., & Steffgen, G. (2021). Lifting the Curtain on Motivation in Exercise: Validation of Two Questionnaires for Physical Education and Leisure Time in French and German. *Journal of Psychoeducational Assessment*, 39(5), 623–639. <https://doi.org/10.1177/07342829211005571>
- International Test Commission. (2018). ITC Guidelines for Translating and Adapting Tests (Second Edition). *International Journal of Testing*, 18(2), 101–134. <https://doi.org/10.1080/15305058.2017.1398166>
- Jankauskiene, R., Urmanavicius, D., & Baceviciene, M. (2022). Associations between Perceived Teacher Autonomy Support, Self-Determined Motivation, Physical Activity Habits and Non-Participation in Physical Education in a Sample of Lithuanian Adolescents. *Behavioral Sciences*, 12(9), 314. <https://doi.org/10.3390/bs12090314>
- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>
- Kaiser, H. F. (1970). A Second Generation Little Jiffy. *Psychometrika*, 35(4), 401–415. <https://doi.org/10.1007/BF02291817>
- Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, 34(1), 111–117. <https://doi.org/10.1177/001316447403400115>
- Kline, P. (2014). *An Easy Guide to Factor Analysis* (0 edn.). Routledge. <https://doi.org/10.4324/9781315788135>
- Kyriazos, T., & Poga, M. (2023). Dealing with Multicollinearity in Factor Analysis: The Problem, Detections, and Solutions. *Open Journal of Statistics*, 13(03), 404–424. <https://doi.org/10.4236/ojs.2023.133020>
- Li, Y., Wen, Z., Hau, K.-T., Yuan, K.-H., & Peng, Y. (2020). Effects of Cross-loadings on Determining the Number of Factors to Retain. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(6), 841–863. <https://doi.org/10.1080/10705511.2020.1745075>
- Lonsdale, C., Sabiston, C. M., Taylor, I. M., & Ntoumanis, N. (2011). Measuring student motivation for physical education: Examining the psychometric properties of the Perceived Locus of Causality Questionnaire and the Situational Motivation Scale. *Psychology of Sport and Exercise*, 12(3), 284–292. <https://doi.org/10.1016/j.psychsport.2010.11.003>
- MacCallum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, 100(1), 107–120. <https://doi.org/10.1037/0033-2909.100.1.107>
- MacCallum, R. C., & Tucker, L. R. (1991). Representing sources of error in the common-factor model: Implications for theory and practice. *Psychological Bulletin*, 109(3), 502–511. <https://doi.org/10.1037/0033-2909.109.3.502>
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111(3), 490–504. <https://doi.org/10.1037/0033-2909.111.3.490>
- McDonald, R. P. (2013). *Test Theory: A Unified Treatment* (0 edn.). Psychology Press. <https://doi.org/10.4324/9781410601087>
- Mullan, E., Markland, D., & Ingledew, D. K. (1997). A graded conceptualisation of self-determination in the regulation of exercise behaviour: Development of a measure using confirmatory factor analytic procedures. *Personality and Individual Differences*, 23(5), 745–752. [https://doi.org/10.1016/s0191-8869\(97\)00107-4](https://doi.org/10.1016/s0191-8869(97)00107-4)
- Muratbayevich, J. P. (2024). PHYSICAL EDUCATION AND ITS IMPACT ON MENTAL HEALTH. *International Journal of Pedagogics*, 4(11), 220–224. <https://doi.org/10.37547/ijp/Volume04Issue11-42>

- Murcia, J. A. M., Coll, D. G.-C., & Garzón, M. C. (2009). Preliminary validation in Spanish of a scale designed to measure motivation in physical education classes: The Perceived Locus of causality (PLOC) scale. *The Spanish Journal of Psychology*, 12(1), 327–337. <https://doi.org/10.1017/S1138741600001724>
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91(3), 328–346. <https://doi.org/10.1037/0033-295x.91.3.328>
- Ntoumanis, N. (2005). A Prospective Study of Participation in Optional School Physical Education Using a Self-Determination Theory Framework. *Journal of Educational Psychology*, 97(3), 444–453. <https://doi.org/10.1037/0022-0663.97.3.444>
- 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(3), 717–728. <https://doi.org/10.1037/a0014696>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3. ed.). McGraw-Hill. [Nachdr.]. https://books.google.co.jp/books/about/Psychometric_Theory_3E.html?id=_6R_f3G58JsC&redir_esc=y
- Radomir, L., & Moisescu, O. I. (2019). Discriminant validity of the customer-based corporate reputation scale: Some causes for concern. *Journal of Product & Brand Management*, 29(4), 457–469. <https://doi.org/10.1108/JPBm-11-2018-2115>
- Raykov, T. (1997). Estimation of Composite Reliability for Congeneric Measures. *Applied Psychological Measurement*, 21(2), 173–184. <https://doi.org/10.1177/01466216970212006>
- Rönkkö, M., & Cho, E. (2022). An Updated Guideline for Assessing Discriminant Validity. *Organizational Research Methods*, 25(1), 6–14. <https://doi.org/10.1177/1094428120968614>
- Roths, A., Lemos, M. S., & Gonçalves, T. (2022). The Influence of Students' Self-Determination and Personal Achievement Goals in Learning and Engagement: A Mediation Model for Traditional and Nontraditional Students. *Education Sciences*, 12(6), 369. <https://doi.org/10.3390/educsci12060369>
- 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(5), 749–761. <https://doi.org/10.1037/0022-3514.57.5.749>
- Ryan, R. M., & Deci, E. L. (Eds.). (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- 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(2), 125–135. <https://doi.org/10.1080/02701367.2012.10599842>
- Salmanov, V., Zeynalov, C., Hasanov, N., Talibova, L., & Salmanova, K. (2025). A Healthy Lifestyle of Schoolchildren and Students as A Result of Systematic Physical Education. *Journal of Neonatal Surgery*, 14(4S), 788–798. <https://doi.org/10.52783/jns.v14.1872>
- Sagr, M., & López-Pernas, S. (Eds.). (2024). *Learning Analytics Methods and Tutorials: A Practical Guide Using R*. Springer Nature. <https://link.springer.com/content/pdf/10.1007/978-3-031-54464-4.pdf>
- Saris, W. E., Satorra, A., & Sorbom, D. (1987). The Detection and Correction of Specification Errors in Structural Equation Models. *Sociological Methodology*, 17, 105. <https://doi.org/10.2307/271030>
- Schunk, D. H., Meece, J. L., & Pintrich, P. R. (2014). *Motivation in education: Theory, research, and applications* (4th ed.). Pearson. https://books.google.co.jp/books/about/Motivation_in_Education.html?id=HFxtMAEACAAJ&redir_esc=y
- Sánchez De Miguel, M., Lizaso, I., Hermosilla, D., Alcover, C., Goudas, M., & Arranz-Freijó, E. (2017). Preliminary validation of the Perceived Locus of Causality scale for academic motivation in the context of university studies (PLOC-U). *British Journal of Educational Psychology*, 87(4), 558–572. <https://doi.org/10.1111/bjep.12164>
- Spearman, C. (1904). General Intelligence, Objectively Determined and Measured. *The American Journal of Psychology*, 15(2), 201. <https://doi.org/10.2307/1412107>
- 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(1), 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(3), 411–433. <https://doi.org/10.1348/000709904x22359>
- Strauss, M. E., & Smith, G. T. (2009). Construct Validity: Advances in Theory and Methodology. *Annual Review of Clinical Psychology*, 5(1), 1–25. <https://doi.org/10.1146/annurev.clinpsy.032408.153639>
- Suleymanova, A. N., & Zangieva, I. K. (2022). Selection of Factor Extraction Methods in Complicated Research Contexts: Practice Recommendations. *Tomsk State University Journal of Philosophy, Sociology and Political Science*, 69, 152–160. <https://doi.org/10.17223/1998863x/69/16>. <https://cyberleninka.ru/article/n/selection-of-factor-extraction-methods-in-complicated-research-contexts-practice-recommendations/viewer>
- Sullivan, J. H., Warkentin, M., & Wallace, L. (2021). So many ways for assessing outliers: What really works and does it matter? *Journal of Business Research*, 132, 530–543. <https://doi.org/10.1016/j.jbusres.2021.03.066>

- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4. ed.). Allyn and Bacon. [Nachdr.]. https://books.google.co.jp/books/about/Using_Multivariate_Statistics.html?id=IVtqAAAAMAAJ&redir_esc=y
- Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2019). *Using multivariate statistics* (Seventh edition). Pearson. <https://www.pearsonhighered.com/assets/preface/0/1/3/4/0134790545.pdf>
- Tavakol, M., & Wetzell, A. (2020). Factor Analysis: A means for theory and instrument development in support of construct validity. *International Journal of Medical Education*, 11, 245–247. <https://doi.org/10.5116/ijme.5f96.0f4a>
- Teixeira, D. S., Monteiro, D., Carraça, E., & Palmeira, A. L. (2018). Translation and validation of the perceived locus of causality questionnaire (PLOCQ) in a sample of portuguese physical education students. *Motriz: Revista de Educação Física*, 24(2), 1–10. <https://doi.org/10.1590/s1980-6574201800020007>
- Tinsley, H. E. A., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, 34(4), 414–424. <https://doi.org/10.1037/0022-0167.34.4.414>
- Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley Pub. Co. <https://www.consoleflare.com/blog/wp-content/uploads/2022/09/Exploratory-Data-Analysis-1977-John-Tukey.pdf>
- Urhahne, D., & Wijnia, L. (2023). Theories of Motivation in Education: An Integrative Framework. *Educational Psychology Review*, 35(2), 45. <https://doi.org/10.1007/s10648-023-09767-9>
- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education. *Educational and Psychological Measurement*, 52(4), 1003–1017. <https://doi.org/10.1177/0013164492052004025>
- Vansteenkiste, M., Lens, W., Elliot, A. J., Soenens, B., & Mouratidis, A. (2014). Moving the Achievement Goal Approach One Step Forward: Toward a Systematic Examination of the Autonomous and Controlled Reasons Underlying Achievement Goals. *Educational Psychologist*, 49(3), 153–174. <https://doi.org/10.1080/00461520.2014.928598>
- Vlachopoulos, S. P., Katartzi, E. S., Kontou, M. G., Moustaka, F. C., & Goudas, M. (2011). The revised perceived locus of causality in physical education scale: Psychometric evaluation among youth. *Psychology of Sport and Exercise*, 12(6), 583–592. <https://doi.org/10.1016/j.psychsport.2011.07.003>
- Wang, J. C. K., Hagger, M., & Liu, W. C. (2009). A Cross-Cultural Validation of Perceived Locus of Causality Scale in Physical Education Context. *Research Quarterly for Exercise and Sport*, 80(2), 313–325. <https://doi.org/10.1080/02701367.2009.10599566>
- Watkins, M. W. (2021). *A step-by-step guide to exploratory factor analysis with SPSS*. Routledge Taylor & Francis Group. <https://doi.org/10.4324/9781003149347>
- Whittaker, T. A., & Schumacker, R. E. (2022). *A beginner's guide to structural equation modeling* (Fifth edition). Routledge. <https://doi.org/10.4324/9781003044017>
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement*, 73(6), 913–934. <https://doi.org/10.1177/0013164413495237>
- Wolf, J., Kleinert, J., Pels, F., & Vogelsang, A. (2022). Translation and validation of an extended German-language version of the Perceived Locus of Causality Questionnaire (PLOCQ-G) in a sample of physical education students. *German Journal of Exercise and Sport Research*, 52(4), 612–625. <https://doi.org/10.1007/s12662-021-00789-5>
- Wuest, D. A., & Bucher, C. A. (2018). Chapter 1 Meaning and Scope. In *Foundations of Physical Education, Exercise Science, and Sport*. McGraw-Hill Education. <https://library.rtc.bt/cgi-bin/koha/opac-detail.pl?biblionumber=9090>
- Yang, M., Viladrich, C., & Cruz, J. (2019). Measuring motivation in physical education among Chinese and Spanish adolescents: Comparing the psychometric properties of perceived locus of causality scale. *Journal of Sport Psychology*, 28(2), 87–96. <https://archives.rpd-online.com/article/view/v28-n2-yang-viladrich-cruz.html>