REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity

www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333

Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

Gender-related differential item functioning and academic predictors of study completion expectations in sports science students

Funcionamiento diferencial del ítem relacionado con el género y predictores académicos de las expectativas de finalización de estudios en estudiantes de ciencias del deporte

Bayu Budi Prakoso[1,*]; Suroto[2]; Siswantoyo[3]; Syukrul Hamdi[3]; Risky Setiawan[3]; Nanang Indriarsa[1]; Dwi Lorry Juniarisca[1]; Ma'rifatun Nashikhah[4]

[1]Physical Education, Health and Recreation Undergraduate Program, Faculty of Sports and Health Sciences, Universitas Negeri Surabaya, Surabaya, Indonesia;
[2]Doctoral Program in Physical Education, Faculty of Sports and Health Sciences, Universitas Negeri Surabaya, Surabaya, Indonesia;
[3]Educational Research and Evaluation Program, Postgraduate School, Universitas Negeri Yogyakarta, Sleman, Indonesia;
[4]Fashion Education Study Program, Faculty of Engineering, Universitas Negeri Surabaya, Surabaya, Indonesia

**Correspondence Author: Bayu Budi Prakoso, bayuprakoso@unesa.ac.id

Editorial Schedule: *Article received: 18/07/2026 Accepted:19/10/2026 Published: 01/11/2026*

To cite this article use the following reference:
Prakoso, B.B.; Suroto; Siswantoyo; Hamdi, S.; Setiawan, R.; Indriarsa, N.; Juniarisca, D.L.; Nashikhah, M. (2026). Gender-related differential item functioning and academic predictors of study completion expectations in sports science students. *Sportis Sci J, 12*(7), 1-23, https://doi.org/10.55040/9t8ebb57

Author contribution: B.B.P.: Conceptualization, methodology, data collection, formal analysis, writing – original draft, writing – review & editing. S. (Suroto): Supervision, writing – review & editing. S. (Siswantoyo): Methodology, writing – review & editing. S.H.: Methodology, formal analysis, writing – review & editing. R.S.: Formal analysis, writing – review & editing. N.I.: Supervision, writing – review & editing. D.L.J.: Writing – review & editing. M.N.: Project administration, writing – review & editing. All authors have read and agreed to the published version of the manuscript.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

Funding: This study was supported by the 2026 Research Grant of Universitas Negeri Surabaya (No. 459/UN38/HK/2026).

Conflict of interest: The authors declare no conflict of interest.

Ethical aspects: This study received ethical approval from STKIP Bangkalan (No. 004/C8/G/I/2026) and was conducted in accordance with the Declaration of Helsinki. All participants provided voluntary informed consent prior to participation.

Abstract
Gender-related differential item functioning and academic predictors of study completion expectations among sports science students remain underexplored despite their theoretical relevance within the Social Cognitive Career Theory (SCCT) framework. This study developed and validated the Study Completion Expectation (SCE) instrument, initially comprising 10 items across two subfactors, among 458 Indonesian sports science students (304 male, 154 female; 116 athletes, 342 non-athletes) through EFA, CFA, and Rasch analysis. Exploratory factor analysis reduced the instrument to a 9-item, two-factor structure with strong psychometric properties (Robust CFI = 0.974; $\alpha = 0.911–0.922$; WLE reliability = 0.774–0.856 for the AOE and PSB subscales and 0.872 for the composite SCE score). A preliminary multi-group CFA, conducted on subsamples below recommended size thresholds, did not indicate non-invariance across gender or athlete status, though this exploratory finding requires replication with larger samples. Rasch theta-based regression confirmed GPA as a significant predictor across all three models ($p < 0.05$; range $p = 0.002–0.022$), consistent with SCCT predictions regarding performance accomplishments. Thesis readiness showed a positive trend toward perceived barriers but did not reach significance ($B = 0.548$, $p = 0.058$). The most notable finding was that female students consistently demonstrated higher outcome-expectation beliefs on three Academic Outcome Expectation items than male students, indicating gender-related differential item functioning (DIF) detectable only through Rasch item-level analysis and invisible at the construct level; a smaller DIF effect was similarly observed on one Perceived Surmountable Barriers item for athlete status. These findings indicate that comparing AOE scores across gender risks obscuring meaningful differences and warrants the development of gender-specific norms.

Keywords: study completion expectation; sports science students; Social Cognitive Career Theory; gender-related differential item functioning (DIF); academic achievement

Resumen
El funcionamiento diferencial del ítem relacionado con el género y los predictores académicos de las expectativas de finalización de estudios en estudiantes de ciencias del deporte permanecen poco explorados, pese a su relevancia teórica en el marco de la Teoría Social Cognitiva de Carrera (SCCT). Este estudio desarrolló y validó el instrumento Study Completion Expectation (SCE), compuesto inicialmente por diez ítems en dos subfactores, en 458 estudiantes indonesios de ciencias del deporte (304 hombres, 154 mujeres; 116 atletas, 342 no

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

atletas), mediante AFE, AFC y análisis de Rasch. El análisis factorial exploratorio redujo el instrumento a una estructura de dos factores con nueve ítems, evidenciando propiedades psicométricas sólidas (CFI robusto = 0,974; α = 0,911–0,922; fiabilidad WLE = 0,774–0,856 para las subescalas AOE y PSB, y 0,872 para la puntuación compuesta de SCE). Un AFC multigrupo preliminar, en submuestras por debajo del tamaño recomendado, no indicó falta de invariancia entre géneros ni entre estatus deportivo, aunque este hallazgo es exploratorio y requiere replicación con muestras mayores. La regresión basada en puntuaciones theta de Rasch confirmó el promedio académico (GPA) como predictor significativo en los tres modelos ($p < ,05$; rango $p = ,002–,022$), acorde con las predicciones de la SCCT sobre los logros de rendimiento. La disposición para la tesis mostró una tendencia positiva hacia las barreras percibidas, sin significación estadística (B = 0,548, p = 0,058). El hallazgo más destacado fue que las mujeres mostraron consistentemente creencias más altas en tres ítems de Academic Outcome Expectation frente a los hombres, indicando un funcionamiento diferencial del ítem (DIF) relacionado con el género, detectable solo mediante análisis Rasch a nivel de ítem e invisible a nivel de constructo; un efecto DIF menor se observó también en un ítem de Perceived Surmountable Barriers según estatus deportivo. Estos hallazgos indican que comparar puntuaciones AOE entre géneros conlleva el riesgo de ocultar diferencias significativas y justifica el desarrollo de normas específicas por género.

Palabras clave: expectativa de finalización de estudios; estudiantes de ciencias del deporte; Teoría Social Cognitiva de Carrera; funcionamiento diferencial del ítem relacionado con el género (DIF); rendimiento académico

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

Introduction

Sports science students face unique challenges that distinguish them from their peers, namely the dual demands of being both academics and active athletes who must manage intensive training schedules, competitions, and travel (Linnér et al., 2019). These demands are not merely physical and competitive but also encompass the simultaneous development of professional competencies as future educators and coaches (Prasetyo et al., 2025; Vidal-Vilaplana et al., 2022). This condition creates significant pressure on students' ability to complete their studies on time (Sum et al., 2017; Tompsett et al., 2023). In Indonesia, this phenomenon is particularly relevant in the context of the dual-career athlete policy mandated by Presidential Regulation No. 86 of 2021 on the National Sports Development Blueprint (DBON, *Desain Besar Olahraga Nasional*), which affirms the need for synergy between athletic achievement and the completion of formal education. Despite this urgency, early detection of students at risk of barriers to study completion remains limited due to the lack of a validated standard instrument in this context (Condello et al., 2019; Lent et al., 2016; Suroto et al., 2024).

Theoretically, study completion expectations can be understood through the Social Cognitive Career Theory (SCCT) framework (Lent et al., 1994), which positions outcome expectations and perceived barriers as two central constructs in the formation of academic and career behavior (Wang et al., 2022). Outcome expectations refer to an individual's beliefs about the consequences of a particular action (Lent et al., 1994) specifically in the academic context, the belief that studying will result in timely graduation. Meanwhile, perceived barriers are obstacles that hinder the achievement of academic goals (Lent & Brown, 2019), such as thesis demands, training loads, and financial constraints (Cosh & Tully, 2015). SCCT asserts that although these two constructs are correlated, they have distinct mechanisms of formation and intervention implications and therefore need to be measured separately (Lent et al., 1994; Lent & Brown, 2019).

Qualitative research by Suroto et al. (2024) has identified central themes related to study completion expectations and barriers among Indonesian sports science students; however, these findings have not yet been operationalized into a validated quantitative instrument (Condello et

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

al., 2019; Lent et al., 2016). Existing international instruments were designed for general populations and do not account for the unique characteristics of sports science students (Linnér et al., 2019; Mateu et al., 2020), necessitating a new instrument that is both contextually relevant and meets rigorous psychometric standards.

This study developed and validated the SCE instrument, specifically designed for sports science students, based on SCCT and the Indonesian sports education context, and distinguished two subfactors (AOE and PSB) with different intervention implications. The study aimed to explore and confirm the factor structure of the SCE instrument; test convergent and discriminant validity and reliability; and examine external relationship-based validity through multiple regression (GPA and thesis readiness) and differential item functioning (DIF) by gender and athlete status.

Methodology

Design and participants

This study employed a quantitative approach with an instrument validation design. Participants were 458 sports science students from a single sports science program at one Indonesian higher education institution (304 male, 154 female; 116 athletes, 342 non-athletes), selected using purposive sampling (Etikan et al., 2016) from active sixth-semester students entering the final phase of their studies. Of these, 186 students had determined their thesis title, while 272 had not. All participants provided voluntary consent through a digital informed consent form. This study received ethical approval from [institution name withheld for blind review] (No. 004/C8/G/I/2026) and was conducted as part of the 2026 Research Grant of [institution name withheld for blind review] (No. 459/UN38/HK/2026).

Instrument

The SCE scale was developed based on the SCCT framework (Lent et al., 1994) and contextual qualitative findings by Suroto et al. (2024), consisting of two subfactors: Academic Outcome Expectation (AOE, 5 items) and Perceived Surmountable Barriers (PSB, 5 items), all using a 5-point Likert scale. Content validity was evaluated by a panel of five multidisciplinary

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

experts (AERA et al., 2014; Lynn, 1986) who assessed the relevance of each item on a four-point scale (Polit & Beck, 2006). Nine of ten items obtained I-CVI = 1.00, with S-CVI/Ave = 0.96, exceeding the minimum standard of 0.90 (Polit & Beck, 2006; Yusoff, 2019). One item (PSB_5, I-CVI = 0.60) was revised to a reverse-scored format based on expert feedback and retained for empirical testing, following recommended practice when an item is problematic in content but remains relevant to the construct (Polit & Beck, 2006; Zamanzadeh et al., 2015).

Data analysis procedure

Data were analyzed using R version 4.4.3 (packages: psych, GPArotation, lavaan, semTools, car, lmtest). Analysis was conducted in three stages. First, data were split into two subsamples (random split-half, n = 229 each, seed = 42). The first subsample was used for EFA using Maximum Likelihood with Varimax rotation; data suitability was tested via KMO and Bartlett's Test of Sphericity; the number of factors was determined based on Kaiser criteria confirmed by Parallel Analysis; items were retained if factor loading ≥ 0.50 and communality ≥ 0.30 (Costello & Osborne, 2005). Second, the second subsample was used for CFA with MLR estimator (lavaan); the model was respecified based on substantively justified modification indices (Byrne, 2013); model evaluation used Robust CFI ≥ 0.95, Robust RMSEA < 0.08, and SRMR < 0.08 (Hu & Bentler, 1999); convergent validity was evaluated through AVE ≥ 0.50 and CR ≥ 0.70 (Fornell & Larcker, 1981); discriminant validity was tested through Fornell-Larcker criterion, √AVE matrix, and HTMT < 0.85 (Henseler et al., 2015). Third, multiple regression (N = 458) tested external relationship-based validity in accordance with the AERA validity framework (Hawkins et al., 2018; Hilkenmeier et al., 2020), with GPA and thesis readiness as predictors; assumptions were tested using Shapiro-Wilk, Durbin-Watson, Breusch-Pagan, and VIF tests.

As a complement to CFA, Rasch Rating Scale Model (RSM) analysis (Andrich, 1978; Chong et al., 2022) was conducted in R using the TAM package, separately for AOE and PSB. Reliability was evaluated through Expected A Posteriori (EAP) and Weighted Likelihood Estimate (WLE) reliability (Warm, 1989); item fit was evaluated through infit and outfit mean-square statistics with productive range 0.5–1.5 (Bond et al., 2020); item-person targeting was

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

evaluated by comparing item thresholds and theta distribution. Differential Item Functioning (DIF) based on gender and athlete status was tested using separate calibration per group and a z-test of item difficulty differences (Bond et al., 2020), with Holm correction for multiple comparisons and a substantive magnitude criterion of $|\Delta b| \geq 0.5$ logits (Bond et al., 2020). WLE theta scores were used as the dependent variable in multiple regression with GPA, thesis readiness, gender, and athlete status as predictors, to examine whether item-level DIF manifested as construct-level score differences.

As a preliminary complement to the item-level DIF analysis, we conducted multi-group CFA (MGCFA) on the CFA subsample (n = 229) to examine measurement invariance across gender and athlete status, using the same two-factor model with correlated error terms. We evaluated the model using a three-step nested comparison (configural, metric, and scalar invariance). Model fit differences were assessed against the $\Delta CFI \leq 0.01$ and $\Delta RMSEA \leq 0.015$ criteria (Putnick & Bornstein, 2016), together with scaled chi-square difference tests. Degrees of freedom for the metric invariance comparison were calculated following the approach recommended for referent-indicator scaling (Klopp & Klößner, 2023). As both gender (n = 156 vs. 73) and athlete-status (n = 166 vs. 63) subgroups fell well below the conventional minimum of n = 200 per group (Hair, Black, et al., 2019), this analysis was treated as exploratory rather than confirmatory.

Results

Descriptive statistics and exploratory factor analysis (EFA)

Data suitability tests yielded KMO = 0.889 and a significant Bartlett's Test of Sphericity ($\chi^2(45) = 1633.57$, $p < 0.001$). Parallel Analysis and the Kaiser criterion indicated a two-factor structure (eigenvalues: 5.43 and 1.79), explaining 66.71% of the total variance after Varimax rotation. Table 1 presents descriptive statistics and factor loadings. AOE items showed high mean scores (M = 4.29–4.47), while PSB items showed more varied means (3.76–3.97). Nine of ten items had factor loadings ≥ 0.70. PSB_5 (communality = 0.119; SD = 1.03) was excluded, yielding a 9-item final instrument: 5 AOE and 4 PSB items. Reliability was excellent: AOE ($\alpha = 0.922$), PSB ($\alpha = 0.914$), and SCE ($\alpha = 0.911$).

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

TABLE 1. EFA Factor Loadings with Descriptive Statistics: SCE Instrument (Varimax, n = 229)

Item	M	SD	F1 (AOE)	F2 (PSB)	Communality	Uniqueness	Decision
ACADEMIC OUTCOME EXPECTATION (AOE)							
AOE_1	4.44	0.69	0.801	0.325	0.748	0.252	Retained
AOE_2	4.29	0.74	0.7	0.405	0.653	0.347	Retained
AOE_3	4.44	0.74	0.794	0.255	0.696	0.304	Retained
AOE_4	4.47	0.68	0.886	0.178	0.817	0.183	Retained
AOE_5	4.43	0.73	0.823	0.13	0.694	0.306	Retained
PERCEIVED SURMOUNTABLE BARRIERS (PSB)							
PSB_1	3.89	0.78	0.206	0.853	0.771	0.229	Retained
PSB_2	3.79	0.82	0.193	0.8	0.678	0.322	Retained
PSB_3	3.76	0.78	0.175	0.862	0.774	0.226	Retained
PSB_4	3.97	0.73	0.318	0.788	0.721	0.279	Retained
PSB_5*	3.4	1.03	0.22	0.265	0.119	0.881	Excluded

NOTE. AOE = Academic Outcome Expectation; PSB = Perceived Surmountable Barriers; M = mean; SD = standard deviation; *Excluded from the final instrument. F1 = factor loading on AOE factor; F2 = factor loading on PSB factor. N = 458 for descriptive statistics; n = 229 for EFA.

Confirmatory factor analysis (CFA)

CFA on the second subsample (n = 229) modified the initial two-factor model by freeing two substantively justified error correlations: AOE_4–AOE_5 (MI = 19.535) and PSB_2–PSB_3 (MI = 12.992) (Byrne, 2013). Robust CFI = 0.974 and SRMR = 0.055 met good fit criteria, indicating a well-fitting model overall. Robust RMSEA = 0.084 [90% CI: 0.056, 0.113], however, exceeded the conventional 0.08 threshold. While this may partly reflect the model's small degrees of freedom, a condition under which RMSEA is known to be less reliable and prone to overrejecting adequately fitting models (Kenny et al., 2015; Marsh et al., 2004), we report this value transparently rather than dismissing it, and interpret overall model fit based on the convergence of multiple indices (CFI, SRMR) rather than RMSEA in isolation. The two-factor model consistently outperformed the one-factor model (Table 2). All factor loadings were significant ($p < 0.001$), ranging from 0.767 to 0.923 (Table 3). Convergent validity was satisfied (AVE ≥ 0.64; CR ≥ 0.877) and discriminant validity confirmed through Fornell-Larcker, √AVE matrix, and HTMT = 0.537 < 0.85 (Henseler et al., 2015).

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

TABLE 2. Comparison of CFA Model Fit Indices by Robust Indices (n = 229)

Index Fit	Cut-off	Model 2 Faktor	Model 1 Faktor
Robust CFI	> 0.95	0.974	0.729
Robust TLI	> 0.95	0.962	0.639
Robust RMSEA	< 0.08	0.084[a]	0.262
90% CI RMSEA	-	[0.056; 0.113]	[0.241; 0.284]
SRMR	< 0.08	0.055	0.162

NOTE. [a] RMSEA slightly exceeds the 0.08 cut-off; see main text for discussion of this value in relation to model degrees of freedom and other fit indices. Error correlations were freed between AOE_4 and AOE_5 (r = 0.404, p = 0.011) and PSB_2 and PSB_3 (r = 0.354, p = 0.086) (Byrne, 2013).

TABLE 3. Standardized Factor Loadings CFA and Construct Validity Indices (n = 229)

Item	λ	SE	z	p
AOE_1	0.901	0.020	45.100	< 0.001
AOE_2	0.811	0.039	20.700	< 0.001
AOE_3	0.820	0.070	11.700	< 0.001
AOE_4	0.923	0.016	58.400	< 0.001
AOE_5	0.776	0.060	12.900	< 0.001
PSB_1	0.785	0.048	16.400	< 0.001
PSB_2	0.822	0.048	17.100	< 0.001
PSB_3	0.767	0.049	15.600	< 0.001
PSB_4	0.824	0.058	14.200	< 0.001
AVE: AOE= 0.719; PSB= 0.640 | CR: AOE= 0.927; PSB= 0.877 | HTMT (AOE.PSB) = 0.537				
Fornell-Larcker Matrix: √AVE_AOE = 0.848; r(AOE.PSB) = 0.495; √AVE_PSB = 0.800				

NOTE. λ = standardized factor loading; SE = standard error; AVE = Average Variance Extracted; CR = Composite Reliability; HTMT = Heterotrait-Monotrait Ratio. All λ significant at p < 0.001.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

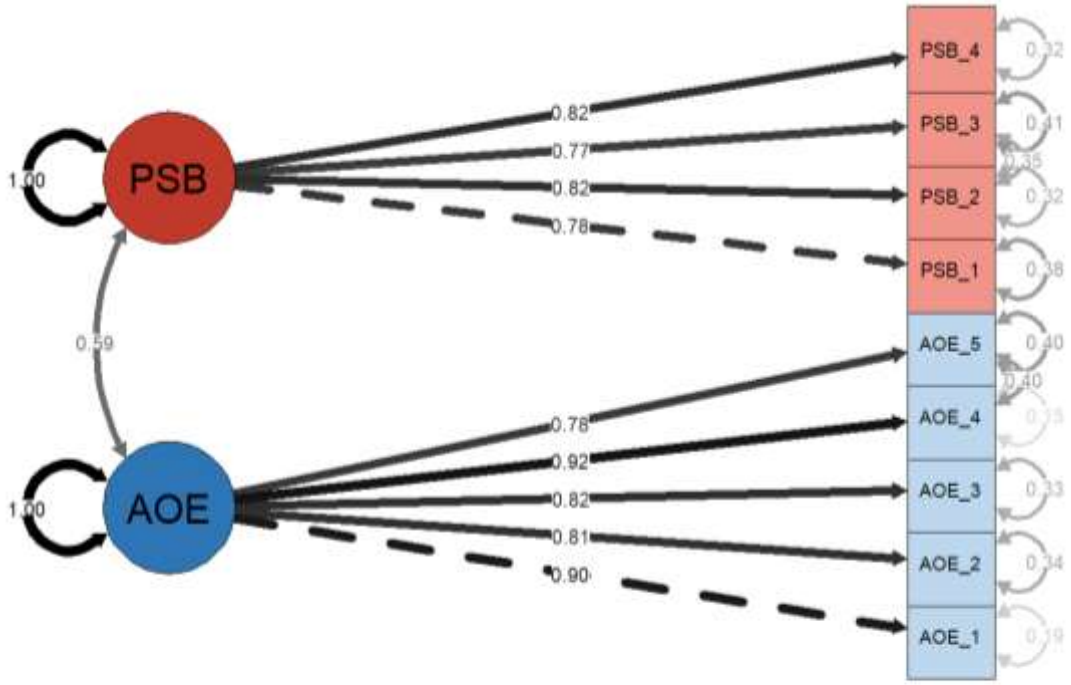


FIGURE 1. Standardized Solution of the Two-Factor CFA Model of the SCE Instrument with Correlated Error Terms (n = 229)

NOTE. Values are standardized estimates. Dashed arrows = marker indicators (loadings fixed for identification). Curved arrows between AOE_4–AOE_5 and PSB_2–PSB_3 = freed error correlations.

Rasch analysis (item response theory)

The SCE instrument was analyzed using the Rasch Rating Scale Model (Andrich, 1978) separately for AOE and PSB (Table 4). EAP and WLE reliability were good to excellent across all constructs (WLE: 0.774–0.872). All items fell within the productive infit/outfit range (0.52–0.92) (Bond et al., 2020). Item difficulty locations were uniformly negative, indicating sensitivity toward respondents with lower expectations, with a ceiling effect limiting discrimination at mid-to-high levels (Bond et al., 2020; Chong et al., 2022).

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

TABLE 4. Rasch Reliability, Item Difficulty and Item Fit: AOE and PSB (N = 458)

Item	b (logit)	Outfit	Infit	Status	EAP Rel.	WLE Rel.
		AOE (5 ITEMS)			0.847	0.774
AOE_1	-5.878	0.641	0.739	Productive fit		
AOE_2	-5.046	-	-	Productive fit		
AOE_3	-5.865	-	-	Productive fit		
AOE_4	-6.057	0.521	0.622	Productive fit		
AOE_5	-5.828	-	-	Productive fit		
		PSB (4 ITEMS)			0.874	0.856
PSB_1	-4.524	0.847	0.914	Productive fit		
PSB_2	-4.054	0.89	0.92	Productive fit		
PSB_3	-3.91	0.775	0.81	Productive fit		
PSB_4	-4.896	0.824	0.902	Productive fit		
		SCE (COMBINED, 9 ITEMS)			0.906	0.872

NOTE. b = item difficulty in logits; EAP Rel. = Expected A Posteriori reliability; WLE Rel. = Weighted Likelihood Estimate reliability. Dash (-) indicates items with outfit and infit within the productive range without special notation. Outfit/infit values < 0.7 reflect highly predictable responses but remain within the acceptable range (Bond et al., 2020).

Differential item functioning (DIF)

Of 18 item × group comparisons, four items showed significant DIF after Holm correction ($|\Delta b| \geq 0.5$ logits): AOE_1, AOE_3, and AOE_5 showed gender DIF, with female students more likely to endorse these items; PSB_1 showed athlete status DIF, with athletes perceiving it as more difficult to overcome than non-athletes. No DIF was found in PSB by gender or in AOE by athlete status (Table 5).

TABLE 5. Items with Significant DIF after Holm Correction (N = 458)

Item	Comparison	b Group 0	b Group 1	Δb	z	p (Holm)
AOE_1	Gender (0=M; 1=F)	-5.719	-6.547	0.828	3.40	0.002
AOE_3	Gender (0=M; 1=F)	-5.665	-6.630	0.965	3.95	<0.001
AOE_5	Gender (0=M; 1=F)	-5.647	-6.547	0.900	3.70	<0.001
PSB_1	Athlete (0=Non; 1=Athlete)	-4.870	-3.944	-0.925	-4.13	<0.001

NOTE: Δb = difference in item difficulty between groups (logits); p (Holm) = p-value after Holm correction for multiple comparisons. Items AOE_2, AOE_4, PSB_2, PSB_3, and PSB_4 showed no significant DIF after correction (p > 0.05) and/or magnitude < 0.5 logits, for either gender or athlete status.

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

Measurement invariance across gender and athlete status

As a preliminary complement to the Rasch-based DIF analysis, multi-group CFA (MGCFA) was conducted on the CFA subsample (n = 229). It is important to note at the outset that subgroup sizes fell substantially short of the minimum sample size conventionally recommended for stable MGCFA estimation (n ≥ 200 per group) (Hair, Black, et al., 2019). Both configural models also produced a near-singular parameter covariance matrix warning, further indicating estimation instability likely attributable to these small subgroup sizes. Given these limitations, interpret the results below as exploratory and provisional rather than confirmatory evidence of invariance.

With this caveat, for gender, configural (Robust CFI = 0.976, RMSEA = 0.082), metric (ΔCFI = 0.001, ΔRMSEA = −0.007), and scalar models (ΔCFI = 0.000, ΔRMSEA = −0.005) did not show meaningful fit deterioration across stages, nominally meeting conventional invariance criteria (Putnick & Bornstein, 2016), with non-significant scaled chi-square difference tests (both $p > 0.50$). For athlete status, the same pattern was observed: configural (Robust CFI = 0.965, RMSEA = 0.101), metric (ΔCFI = 0.001, ΔRMSEA = −0.008), and scalar models (ΔCFI = 0.001, ΔRMSEA = −0.006) similarly did not indicate meaningful non-invariance (both χ^2 difference tests $p > 0.55$). However, given the sample size limitations described above, these findings should not be treated as definitive evidence of full scalar invariance, and formal replication with adequately powered subgroups is needed before this conclusion can be considered established.

External relationship-based validity

GPA was the only significant predictor across all three models ($p < 0.05$; range $p = 0.002–0.022$), consistent with SCCT predictions regarding the role of performance accomplishments (Table 6). Thesis readiness showed a positive trend toward PSB (B = 0.548, $p = 0.058$) but was non-significant for AOE and SCE. Gender and athlete status were non-significant across all models ($p > 0.19$), confirming no construct-level differences consistent with item-level DIF findings. All regression assumptions were satisfied (max VIF = 1.071; Durbin-Watson = 1.90–1.93; Breusch-Pagan $p > 0.35$).

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

TABLE 6. Multiple Regression with WLE Theta Scores (N = 458)

Predictor	B (p) - theta AOE	B (p) - theta PSB	B (p) - theta SCE
GPA	1.432 (0.002**)	1.253 (0.022*)	1.263 (0.003**)
D_Ready	-0.214 (0.381 ns)	0.548 (0.058 ns)	0.139 (0.529 ns)
D_Gender	0.109 (0.671 ns)	-0.400 (0.191 ns)	-0.162 (0.489 ns)
D_Athlete	-0.072 (0.796 ns)	-0.177 (0.588 ns)	-0.119 (0.635 ns)
R^2	0.025	0.024	0.023
F(4.453)	2.924 (p=0.021*)	2.727 (p=0.029*)	2.700 (p=0.030*)

NOTE. B = unstandardized regression coefficient; D_Ready = thesis readiness (0 = not ready, 1 = ready); D_Gender = gender (0 = male, 1 = female); D_Athlete = athlete status (0 = non-athlete, 1 = athlete); ns = non-significant; * $p < 0.05$; ** $p < 0.01$.

Discussion

The findings confirm that the two-factor structure (AOE and PSB) is fully consistent with the SCCT framework (Lent et al., 1994), distinguishing outcome expectations from perceived barriers as constructs with different antecedents and intervention implications that must be measured separately (Lent & Brown, 2019; Wang et al., 2022). In the context of sports science students, this distinction is particularly relevant, as a student-athlete may simultaneously hold high outcome expectations yet perceive training barriers as difficult to overcome.

The exclusion of PSB_5 was based on the convergence of three evidence layers: expert judgment (I-CVI = 0.60, below the minimum standard of 0.78), empirical data (communality = 0.119, well below the 0.30 threshold), and psychometric evidence that reverse-worded items are prone to interpretive confusion, particularly among samples unfamiliar with negatively worded formats (Almanasreh et al., 2019; Costello & Osborne, 2005; Hair, Risher, et al., 2019; Lynn, 1986; van Sonderen et al., 2013). This convergence provides robust justification for exclusion (Boateng et al., 2018).

Once the final nine-item instrument was established, further testing confirmed adequate psychometric properties. The freed error correlations between AOE_4–AOE_5 and PSB_2–PSB_3 were substantively justified by content similarity between adjacent items (Byrne, 2013). Convergent validity was satisfied (AVE ≥ 0.64; CR ≥ 0.877), exceeding recommended

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

thresholds (Cheung et al., 2024; Fornell & Larcker, 1981), and discriminant validity was confirmed simultaneously through Fornell-Larcker, √AVE matrix, and HTMT = 0.537, representing best practice as no single approach is sufficient alone (Henseler et al., 2015; Rönkkö & Cho, 2022). Reliability was excellent across all levels (α = 0.911–0.922). Rasch analysis reinforced these psychometric findings, showing good to excellent reliability (WLE: 0.774–0.872) and productive item fit for all items (Bond et al., 2020), with a ceiling effect indicating the need for higher-difficulty items in future instrument development (Chong et al., 2022).

Beyond strong psychometric properties, the regression findings provide external relationship-based validity evidence, although effect sizes were modest (R^2 ranging from 0.023 to 0.025 across models; see Limitations). All three regression models consistently identified GPA as the only significant predictor of AOE, PSB, and SCE. What supports the theoretically expected link with SCCT here is the direction and statistical significance of this relationship, not the magnitude of variance explained. This finding is fully consistent with SCCT: strong academic achievement reinforces outcome expectations through the performance accomplishments mechanism and increases the perception that barriers are surmountable (Lent et al., 1994; Lent & Brown, 2019). In the context of sports science students, GPA also reflects time management capacity and perseverance in navigating the dual role of student and athlete (Şahin et al., 2018; Suroto et al., 2024).

The most compelling finding was the differential pattern of thesis readiness toward AOE and PSB. Thesis readiness showed a positive trend toward PSB (B = 0.548) approaching the conventional significance threshold (p = 0.058), but not toward AOE (B = -0.209, p = 0.389). This can be understood through the differing characteristics of the two constructs within the SCCT framework: outcome expectations are formed from an accumulation of long-term academic experiences and are relatively stable, whereas perceived barriers are more responsive to contextual conditions that are situational and proximal (Lent et al., 2000; Lent & Brown, 2019). When students have clarity about their thesis topic, barriers that were previously abstract become more defined and feel more surmountable (Cosh & Tully, 2015). This differential directional pattern is consistent with the assumption that AOE and PSB are not identical

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain

constructs, although the marginal significance of PSB ($p = 0.058$) means this evidence is tentative and requires replication: if AOE and PSB were identical constructs, both would respond to the same predictors in the same way (Rönkkö & Cho, 2022).

DIF testing based on gender and athlete status revealed item-level non-invariance: three AOE items showed moderate-to-large gender DIF ($|\Delta b| = 0.83–0.97$ logits, after Holm correction) in a consistent direction, with female students more likely to endorse AOE items than male students, and one PSB item showed athlete status DIF (PSB_1; $|\Delta b| = 0.925$ logits), where athletes perceived the item as a more difficult barrier to overcome than non-athletes. Both findings are non-canceling DIF and are substantively relevant. These findings indicate that the AOE subscale is not yet fully measurement-equivalent between male and female students (Putnick & Bornstein, 2016), and therefore comparisons of AOE and SCE scores across gender should be made cautiously and ideally using gender-specific norms.

For athlete status, PSB showed DIF on one item (PSB_1) while AOE showed no DIF; for gender, the pattern was reversed: AOE showed DIF on three items while PSB showed none. This asymmetry indicates that the two subscales are differentially sensitive to different grouping variables, meaning PSB scores can be compared across gender with more confidence than AOE scores, and AOE scores can be compared across athlete status with more confidence than PSB scores. This heterogeneity in DIF location is itself informative given the diverse nature of sports science student populations (Henriksen et al., 2010; Linnér et al., 2019; Lupo et al., 2016).

Although the MGCFA results did not indicate meaningful non-invariance, the pronounced sample size limitations described above mean these results should be treated as preliminary rather than confirmatory. Rather than reading this as evidence that stands alongside the item-level DIF findings on equal footing, we present it as an indicative, complementary signal: the absence of detected non-invariance at the construct level is consistent with, but does not independently confirm, the localized nature of the item-level DIF findings (three of five AOE items for gender; one of four PSB items for athlete status). This pattern illustrates the value of combining IRT-based DIF analysis, which does not require the same large per-group samples, with CFA-based invariance testing (Putnick & Bornstein, 2016); however, in the

REVISTA INTERNACIONAL SPORTIS Scientific Journal of School Sport, Physical Education and Psychomotricity
www.sportisjournal.com | v.arufe@udc.es | ISSN 2386-8333
Edited by Campus Educa Sportis. A Coruña. Spain
Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

present study the latter should be regarded as a provisional check rather than a definitive validation, pending replication with larger and more balanced subgroup samples.

The universality of the instrument aligns with its practical flexibility. The SCE instrument can be used in three ways: a separate AOE score to measure outcome expectations, a separate PSB score to measure perceived barriers, and an SCE composite score for rapid screening. Theoretically, the SCE composite score can be understood within a second-order factor model framework in which AOE and PSB represent manifestations of a more abstract latent construct, namely general study completion expectation (Cavicchia & Vichi, 2022). In the Indonesian national policy context, the *Desain Besar Olahraga Nasional* (DBON, National Blueprint for Sports Development, Presidential Regulation No. 86/2021) mandates the integration of athletic development and formal education completion for student-athletes. Within this framework, the SCE instrument, once validated across a broader range of institutions, may offer a useful early detection tool for sports science students at risk of study completion barriers, potentially enabling academic advisors and program coordinators to provide proactive and targeted support before students reach a critical point of dropout risk (Quinaud et al., 2022; Storm & Eske, 2022). Because this validation was conducted at a single institution, this application should be considered a direction for future work rather than an immediate implication of the current findings.

Nevertheless, several limitations must be acknowledged to maintain the integrity of the findings. The sample was drawn from a single sports science program at one institution, so generalization to other sports science programs in Indonesia should be approached cautiously; replication with samples from diverse institutions and regions is strongly recommended to strengthen the external validity of the instrument (Condello et al., 2019). Furthermore, although a preliminary MGCFA nominally supported measurement invariance across both gender and athlete status, both subgroup comparisons fell far short of recommended sample size guidelines, and this result should not be treated as confirmatory pending replication with adequately powered subgroups (Hair, Black, et al., 2019; Makransky et al., 2015; Monteiro et al., 2019; Putnick & Bornstein, 2016). Robust RMSEA (0.084) also exceeds the conventional 0.080 threshold. This may partly reflect the model's small degrees of freedom (Kenny et al., 2015;

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

Marsh et al., 2004) and we report it here as a genuine limitation rather than explain it away. Finally, the small R^2 values (ranging from 2.3% to 2.5%) indicate that many other variables remain unexplored, such as academic self-efficacy, social support, and intrinsic motivation, which in previous research have been shown to contribute to students' academic achievement and study completion (Lent & Brown, 2019).

Conclusions

This study developed and validated the 9-item SCE scale for Indonesian sports science students grounded in SCCT, demonstrating strong psychometric properties: content validity (S-CVI/Ave = 0.96), confirmed two-factor structure (Robust CFI = 0.974, SRMR = 0.055), convergent validity (AVE ≥ 0.64), discriminant validity (HTMT = 0.537), and reliability across classical and Rasch approaches (α = 0.911–0.922; WLE 0.774–0.872). GPA was the only consistent significant predictor of all three constructs, consistent with SCCT predictions. The differential pattern between the trait-like AOE and the situationally responsive PSB suggests distinct intervention approaches are needed in academic advising. Gender-related DIF on three AOE items, detectable only through Rasch item-level analysis and invisible at the construct level, indicates that cross-gender AOE score comparisons require caution and gender-specific norms; a smaller DIF effect was similarly observed on one PSB item for athlete status. A preliminary multi-group CFA conducted on small subsamples nominally supported measurement invariance across both gender and athlete status at the construct level, complementing rather than contradicting the item-level DIF findings, though this result should be treated as exploratory pending replication with larger and more balanced subgroups. This pattern highlights the added value of complementing classical test theory with IRT-based methods in instrument validation. Pending multi-institutional replication with adequately powered subgroups, the SCE scale may hold potential for integration into academic early detection systems in sports science programs, particularly within Indonesia's dual-career athlete policy context; at this stage, however, its use should remain confined to research settings rather than operational deployment.

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

References

AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. American Educational Research Association. https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, *15*(2), 214–221. https://doi.org/10.1016/j.sapharm.2018.03.066

Andrich, D. (1978). A Rating Formulation for Ordered Response Categories. *Psychometrika*, *43*(4), 561–573. https://doi.org/10.1007/bf02293814

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health*, *6*, 366616. https://doi.org/10.3389/fpubh.2018.00149

Bond, T. G., Yan, Z., & Heene, M. (2020). Applying the rasch model: Fundamental measurement in the human sciences. In *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Taylor and Francis. https://doi.org/10.4324/9780429030499

Byrne, B. M. (2013). Structural equation modeling with amos: basic concepts, applications, and programming. In *Structural Equation Modeling With AMOS*. Psychology Press. https://doi.org/10.4324/9781410600219

Cavicchia, C., & Vichi, M. (2022). Second-Order Disjoint Factor Analysis. *Psychometrika*, *87*(1), 289–309. https://doi.org/10.1007/S11336-021-09799-6

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, *41*(2), 745–783. https://doi.org/10.1007/s10490-023-09871-y

Chong, J., Mokshein, S. E., & Mustapha, R. (2022). Applying the rasch rating scale model (rsm) to investigate the rating scales function in survey research instrument. *Jurnal Cakrawala Pendidikan*, *41*(1), 97–111. https://doi.org/10.21831/CP.V41I1.39130

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

Condello, G., Capranica, L., Doupona, M., Varga, K., & Burk, V. (2019). Dual-career through the elite university student-athletes' lenses: The international FISU-EAS survey. *PLOS ONE*, *14*(10), e0223278. https://doi.org/10.1371/journal.pone.0223278

Cosh, S., & Tully, P. J. (2015). Stressors, coping, and support mechanisms for student athletes combining elite sport and tertiary education: implications for practice. *The Sport Psychologist*, *29*(2), 120–133. https://doi.org/10.1123/tsp.2014-0102

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, and Evaluation*, *10*(1). https://doi.org/10.7275/jyj1-4868

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39. https://doi.org/10.2307/3151312

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning, EMEA.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hawkins, M., Elsworth, G. R., & Osborne, R. H. (2018). Application of validity theory and methodology to patient-reported outcome measures (PROMs): building an argument for validity. *Quality of Life Research 2018 27:7*, *27*(7), 1695–1710. https://doi.org/10.1007/S11136-018-1815-6

Henriksen, K., Stambulova, N., & Roessler, K. K. (2010). Successful talent development in track and field: Considering the role of environment. *Scandinavian Journal of Medicine and Science in Sports*, *20*(SUPPL. 2), 122–132. https://doi.org/10.1111/j.1600-0838.2010.01187.x

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

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

Hilkenmeier, F., Bohndick, C., Bohndick, T., & Hilkenmeier, J. (2020). Assessing distinctiveness in multidimensional instruments without access to raw data – a manifest fornell-larcker criterion. *Frontiers in Psychology*, *11*, 504969. https://doi.org/10.3389/fpsyg.2020.00223

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis : Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods and Research*, *44*(3), 486–507. https://doi.org/10.1177/0049124114543236

Klopp, E., & Klößner, S. (2023). Scaling metric measurement invariance models. *Methodology*, *19*(3), 192–227. https://doi.org/10.5964/meth.10177

Lent, R. W., & Brown, S. D. (2019). Social cognitive career theory at 25: Empirical status of the interest, choice, and performance models. *Journal of Vocational Behavior*, *115*, 103316. https://doi.org/10.1016/j.jvb.2019.06.004

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*(1), 79–122. https://doi.org/10.1006/jvbe.1994.1027

Lent, R. W., Brown, S. D., & Hackett, G. (2000). Contextual supports and barriers to career choice: A social cognitive analysis. *Journal of Counseling Psychology*, *47*(1), 36–49. https://doi.org/10.1037/0022-0167.47.1.36

Lent, R. W., Taveira, M. do C., Figuera, P., Dorio, I., Faria, S., & Gonçalves, A. M. (2016). Test of the social cognitive model of well-being in spanish college students. *Journal of Career Assessment*, *25*(1), 135–143. https://doi.org/10.1177/1069072716657

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

Linnér, L., Stambulova, N. B., Lindahl, K., & Wylleman, P. (2019). Swedish university student-athletes' dual career scenarios and competences. *International Journal of Sport and Exercise Psychology*, *23*(3), xxxvii–Lii. https://doi.org/10.1080/1612197X.2019.1611898

Lupo, C., Mosso, C. O., Guidotti, F., Cugliari, G., Pizzigalli, L., & Rainoldi, A. (2016). Motivation toward dual career of Italian student-athletes enrolled in different university paths. *Sport Sciences for Health 2016 13:3*, *13*(3), 485–494. https://doi.org/10.1007/S11332-016-0327-4

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382–386. https://doi.org/10.1097/00006199-198611000-00017

Makransky, G., Rogers, M. E., & Creed, P. A. (2015). Analysis of the construct validity and measurement invariance of the career decision self-efficacy scale. *Journal of Career Assessment*, *23*(4), 645–660. https://doi.org/10.1177/1069072714553555

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing hu and bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Mateu, P., Inglés, E., Torregrossa, M., Marques, R. F. R., Stambulova, N., & Vilanova, A. (2020). Living life through sport: the transition of elite spanish student-athletes to a university degree in physical activity and sports sciences. *Frontiers in Psychology*, *11*, 539730. https://doi.org/10.3389/fpsyg.2020.01367

Monteiro, D., Teixeira, D. S., Vitorino, A., Moutão, J., Rodrigues, F., Machado, S., & Cid, L. (2019). Behavioral regulation sport questionnaire: gender and sport invariance in portuguese athletes. *Perceptual and Motor Skills*, *126*(2), 323–341. https://doi.org/10.1177/0031512519825700

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

Prasetyo, M. T., Fariz, M., Putra, P., & Babingga, H. (2025). Cardiorespiratory capacity of papuan sports students: what is the cardiorespiratory capacity of papuan sports students? *Bima Loka: Journal of Physical Education*, *6*(1), 11–20. https://doi.org/10.26740/bimaloka.v6i1.48477

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Quinaud, R. T., Capranica, L., Doupona, M., & Guidotti, F. (2022). The holistic development of talented sportspersons through dual-career. *Frontiers in Sports and Active Living*, *4*. https://doi.org/10.3389/fspor.2022.929981

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

Şahin, E., Çekin, R., & Özçelik, İ. Y. (2018). Predictors of academic achievement among physical education and sports undergraduate students. *Sports*, 6(1), 8. https://doi.org/10.3390/sports6010008

Storm, R. K., & Eske, M. (2022). Dual careers and academic achievements: does elite sport make a difference? *Sport, Education and Society*, *27*(6), 747–760. https://doi.org/10.1080/13573322.2021.1919070

Sum, R. K. W., Tsai, H. H., Ching Ha, A. S., Cheng, C. F., Wang, F. J., & Li, M. (2017). Social-ecological determinants of elite student athletes' dual career development in hong kong and taiwan. *SAGE Open*, *7*(2). https://doi.org/10.1177/2158244017707798

Suroto, S., Kusnanik, N. W., Pradana, F. G. A., Prakoso, B. B., Bafirman, B., Barlian, E., Usman, A., Arimbi, A., Sumaryanti, S., & Wijaya, M. A. (2024). Education profiling for east java student-athletes: planning, process, assessment, and expectations. *JOSSAE (Journal of Sport Science and Education)*, *9*(1), 36–45. https://doi.org/10.26740/jossae.v9n1.p36-45

Original article. Gender-related differential item functioning and academic predictors of study completion expectations in sports science students.
Vol. 12, n.º 7; p. 1-23, November 2026. https://doi.org/10.55040/9t8ebb57

Tompsett, J., Mcdossi, O., & Roscigno, V. J. (2023). Collegiate sports participation, academic achievement, and bachelor's degree completion. *Sociological Forum*, *38*(4), 1266–1287. https://doi.org/10.1111/socf.12967

van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of Reverse Wording of Questionnaire Items: Let's Learn from Cows in the Rain. *PLOS ONE*, *8*(7), e68967. https://doi.org/10.1371/journal.pone.0068967

Vidal-Vilaplana, A., Valantine, I., Staskeviciute-Butiene, I., González-Serrano, M. H., Capranica, L., & Calabuig, F. (2022). Combining sport and academic career: Exploring the current state of student-athletes' dual career research field. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *31*, 100399. https://doi.org/10.1016/j.jhlste.2022.100399

Wang, D., Liu, X., & Deng, H. (2022). The perspectives of social cognitive career theory approach in current times. *Frontiers in Psychology*, *13*, 1023994. https://doi.org/10.3389/fpsyg.2022.1023994

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*(3), 427–450. https://doi.org/10.1007/bf02294627

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and implementation content validity study: development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences*, *4*(2), 165–178. https://doi.org/10.15171/jcs.2015.017