International Journal of Educational Research 141 (2027) 103111

Contents lists available at ScienceDirect

International Journal of Educational Research

journal homepage: www.elsevier.com/locate/ijedures

When engagement matters: Agentic engagement, expressive behavior, and academic performance in higher education

Uriel Eduardo Torres Castro

Faculty of Education, Universidad de La Sabana, Chía, Cundinamarca, Colombia

ARTICLE INFO

Keywords:
Agentic engagement
Student agency
Academic performance
Grade point average
Student voice
Measurement invariance
Colombian higher education

ABSTRACT

This study examines whether agentic engagement (AE), understood as students' observable contributions to instructional interaction, can be measured reliably and comparably across women and men and how it relates to academic performance in higher education. Using data from 1713 final-year undergraduate students from six Colombian universities, the study assesses the psychometric properties of the Agentic Engagement Scale (AES) and analyzes its association with grade point average (GPA), which is used as an institutional indicator of academic performance, controlling for gender and academic discipline. The results support the reliability, unidimensional structure, and measurement invariance of the AES across women and men, indicating that AE scores can be meaningfully compared between these gender groups. Regression analyses revealed a positive but modest association between AE and GPA. At the item level, however, the association was identified for AES2, which captures students' expression of preferences and opinions. This pattern indicates that the AES total score should not be interpreted as showing a uniform association with GPA. Instead, the AE–GPA relationship depends on specific measured instructional behaviors through which student initiative becomes visible in teaching interactions. By showing that the AE–GPA association was concentrated in students' expression of preferences and opinions, this study supports a behavior-specific and context-sensitive interpretation of AE in higher education.

1. Introduction

Student engagement research helps interpret why academic outcomes differ across students and settings. Evidence links engagement to achievement and grade point average (GPA), but the association varies across dimensions, measures, and contexts (Carini et al., 2006; Lei et al., 2018; Wong et al., 2024; Zhoc et al., 2019). GPA offers a comparable institutional record of assessed performance. It is also an incomplete basis for interpretation because it records outcomes without showing how student actions, instructional opportunities, and assessment practices relate to those outcomes (Pellegrino et al., 2001; York et al., 2015). The same GPA can reflect different paths through a course or program. The analytical problem is to identify which forms of engagement are most closely associated with assessed performance.

Engagement has long been treated as a multidimensional construct. Behavioral, emotional, cognitive, social, and agentic engagement describe different ways in which students participate in, invest in, relate to, and contribute to educational activity (Fredricks et al., 2004; Reeve & Tseng, 2011; Wong et al., 2024). Meta-analyses indicate that these dimensions show different associations with academic achievement (Lei et al., 2018; Wong et al., 2024). Psychometric work on integrated measures supports the use

E-mail address: urieltoca@unisabana.edu.co.

https://doi.org/10.1016/j.ijer.2026.103111
Received 4 April 2026; Received in revised form 26 August 2026; Accepted 5 September 2026
Available online 10 September 2026

of total engagement scores when warranted, but these scores still need to be interpreted in light of the dimensions and indicators that compose them (Moreira et al., 2020). Studies of engagement indicators make this issue more concrete because specific indicators show different strengths and patterns of association with academic performance (Ogunsakin et al., 2021).

Agentic engagement (AE) narrows the analysis to students' instructional contributions. AE refers to students' constructive contributions to the flow of instruction, including asking questions, expressing preferences and opinions, communicating needs, and making interests visible to the teacher (Reeve, 2013; Reeve & Tseng, 2011). This definition separates AE from student agency. Agency refers to a broader capacity for intentional action across educational contexts, whereas AE captures observable instructional behaviors through which students try to shape teaching and learning conditions (Reeve, 2013; Reeve et al., 2020). Because these contributions are directed to the teacher, they can prompt feedback, clarification, and instructional adjustment. AE also differs from behavioral, emotional, and cognitive engagement because it captures student-initiated actions that can alter instructional interaction and elicit teacher support (Jang et al., 2024; Reeve et al., 2022; Reeve & Tseng, 2011).

Establishing the reliability, factorial structure, and measurement equivalence of the Agentic Engagement Scale (AES) is necessary before interpreting comparisons across women and men (van de Vijver & Tanzer, 2004). Beyond measurement quality, the key issue is whether the AE–GPA association appears at the scale level or in particular agentic behaviors. A scale may be coherent while its items retain different associations with assessed performance. This is plausible for the AES because its items represent distinct instructional actions: communicating needs, expressing preferences and opinions, asking for support, asking questions, and sharing interests (Reeve, 2013). A total score can show whether AE is associated with GPA. Item-level evidence shows which instructional actions are individually associated with GPA. The distinction keeps the empirical claim at the level of measured behavior.

These questions are especially relevant in Latin American higher education, where expanded access has coexisted with inequality and institutional stratification (Anzelin et al., 2025; Labraña & Brunner, 2022). Students in the region form a heterogeneous population whose participation reflects socioeconomic background, cultural capital, prior schooling, and institutional location (Parra Sandoval, 2021). In Colombia, financial constraints, private-sector stratification, and regional differences shape access, persistence, performance, and attrition (Martínez et al., 2015; Melguizo et al., 2016; Páez & Teelken, 2016). Policy reforms promote student-centered learning and outcome-based quality assurance, but lecture-centered traditions and uneven instructional conditions still shape the opportunities students have to ask questions, express preferences, communicate needs, and influence instruction (Torres Castro & Pineda-Baéz, 2026). These conditions make Colombia a relevant setting for examining AE as a form of participation whose expression depends partly on the instructional opportunities available to students.

This study pursues two interrelated objectives. The first assesses the reliability, factorial structure, and measurement invariance of the AES across women and men to determine whether AE can be measured reliably and equivalently in Colombian higher education. The second analyzes whether AE relates to GPA through the AES total score or through specific agentic behaviors, while accounting for gender and academic discipline. The study distinguishes between scale-level and item-level associations, allowing the analysis to determine whether the relationship with GPA is distributed across AE or concentrated in particular instructional behaviors. GPA is treated as an indicator of assessed performance rather than as a direct measure of learning.

2. Agentic engagement and academic performance in higher education: measurement and behavioral specificity

Academic performance refers here to institutionally evaluated performance. GPA summarizes assessed performance across courses. Academic success includes achievement, learning, persistence, satisfaction, skill acquisition, and postcollege outcomes; GPA captures only the graded achievement component (York et al., 2015). University GPA also reflects prior preparation, motivation, self-regulation, personality, socioeconomic conditions, institutional environments, and educational experiences (Pascarella & Terenzini, 2005; Richardson et al., 2012; Rodríguez-Hernández et al., 2020). As an aggregate assessment record, GPA combines performances produced through different learning goals, assessment formats, grading standards, and disciplinary expectations (Koper et al., 2015; Steenman et al., 2016; Tomkin & West, 2022; Westrick, 2017). Assessment theory treats such indicators as evidence for inference, not as direct evidence of learning (Pellegrino et al., 2001). Accordingly, associations between engagement and GPA are interpreted here as associations with assessed performance.

Student engagement helps locate the processes through which involvement is associated with academic performance. Foundational work defined engagement through students' time and effort in educationally purposeful activities and the institutional conditions that support such involvement (Kuh, 2003; Kuh et al., 2008). Recent evidence shows that engagement is multidimensional and that its relationship with achievement and GPA varies by dimension, measure, and context (Carini et al., 2006; Fredricks et al., 2004; Lei et al., 2018; Wong et al., 2024; Zhoc et al., 2019). Discipline, institutional setting, student characteristics, and educational environment further shape this relationship (Leach, 2016; Pascarella & Terenzini, 2005).

This variation is theoretically important because behavioral, emotional, cognitive, social, and agentic dimensions do not relate to assessed performance in the same way; they capture participation, affect, strategy use, relation, and contribution to instruction (Fredricks et al., 2004; Lei et al., 2018; Wong et al., 2024). Integrated engagement scores can be useful when supported psychometrically, but their interpretation still depends on the dimensions and indicators represented in the score (Moreira et al., 2020). Indicator-level evidence reinforces this point because specific engagement indicators can relate differently to academic performance (Ogunsakin et al., 2021).

AE operates as a distinct dimension within student engagement, which traditionally comprises behavioral, emotional, and cognitive components (Fredricks et al., 2004). It adds a student-initiated component to this multidimensional view by identifying how students contribute to instruction rather than only how they respond to it. Behavioral engagement captures effort, attention, persistence, and compliance; emotional engagement refers to interest, enjoyment, boredom, or anxiety; and cognitive engagement describes

strategic investment, self-regulation, and deep processing (Fredricks et al., 2004; Reeve & Tseng, 2011). AE differs because it captures proactive and constructive attempts to shape instruction, including questions, preferences, opinions, needs, requests for support, and expressed interests (Reeve, 2013; Reeve & Shin, 2020). This distinction helps identify student-initiated contributions that general participation or compliance may obscure (Kingsford-Smith et al., 2026; Reeve et al., 2020).

Student agency and AE are related but not equivalent. From social–cognitive and self-regulation perspectives, agency is defined as the capacity to influence actions and environments through intentionality, self-regulation, reflection, and perceived efficacy (Bandura, 2001, 2006; Zimmerman, 2000). Recent higher education research treats agency as situated, relational, and institutionally conditioned, shaped by students' histories, identities, relationships, rules, and available opportunities (Klemenčič, 2023; Mameli et al., 2023; Stenalt & Lassesen, 2022). AE has a narrower scope. It captures a situated instructional expression of agency through observable and reportable classroom behaviors (Reeve, 2013; Reeve & Jang, 2022). The AES measures specific student contributions to instruction, not agency, in its full theoretical breadth (Torres Castro & Pineda-Baéz, 2026).

Self-regulated learning theory helps interpret why agentic behaviors can become relevant to assessed performance. Self-regulated learning involves goal setting, strategic action, monitoring, and adjustment over time (Pintrich, 2000; Zimmerman, 2000). These processes often become visible in instruction when students identify gaps in understanding, seek clarification, request resources, or adjust their approach after feedback. AE captures the communicative side of this regulation in classroom interactions. Students ask for clarification, request resources, express preferences, and communicate interests, which can create opportunities to adjust strategies with instructional support (Reeve, 2013; Reeve et al., 2020). AE is not an additional trait alongside self-regulation. It offers an observable channel through which students can make regulatory needs available for instructional response. Agentic behaviors may therefore be relevant to assessed performance without implying a causal mechanism.

AE operates at the instructional interface because it links students' expressed needs, questions, preferences, and suggestions with teaching decisions. Kahu and Nelson (2018) described engagement in the educational interface, where student and institutional factors interact through teaching, support, and learning conditions. Within that space, AE specifies the student-initiated side of the exchange. Students make instructional needs visible, and teachers may respond by adjusting explanations, pacing, resources, tasks, or classroom structures (Reeve, 2013; Reeve & Jang, 2022; Reeve & Shin, 2020). Longitudinal and week-to-week evidence shows reciprocal relationships between AE and autonomy support, while teacher control or deficit-based beliefs can constrain this process (Jang et al., 2024; Michou et al., 2023; Patall et al., 2022; Zambrano et al., 2025). The educational significance of agentic behaviors depends on the instructional conditions that invite, recognize, or respond to them.

Both the expression of AE and its relationship with academic performance are context dependent. Cultural norms, institutional structures, disciplinary practices, and gendered expectations can shape which forms of student initiative are expressed, recognized, or rewarded. When student initiative is shaped by context, group comparisons require evidence that the instrument represents the same construct across the groups being compared (van de Vijver & Tanzer, 2004). Measurement equivalence matters because score variation is interpretable only when it can be linked to the underlying construct rather than to bias in how the instrument functions. For comparisons across women and men, measurement invariance supports interpreting observed score differences as substantive rather than as artifacts of differential measurement. In underrepresented higher education contexts, this requirement is especially important because engagement constructs often travel across languages, cultures, and institutional traditions.

These considerations support two levels of interpretation for AE. At the scale level, the AES represents a coherent construct only if its reliability, factorial structure, and invariance support that interpretation. At the behavioral level, the five items represent different instructional contributions, so their associations with GPA need not mirror the total score. A coherent AE scale can coexist with behavior-specific associations because expressing preferences, communicating needs, asking for support, asking questions, and sharing interests may enter instruction in different ways. The study's empirical interpretation rests on this distinction: the total AES score addresses AE as a construct, while item-level evidence addresses which agentic behaviors are most closely associated with assessed performance.

3. Methods

3.1. Study design

This study used a quantitative cross-sectional design to examine AE among Colombian university students and its association with academic performance. The design aligned with the study aims because it allowed associations between constructs to be estimated and measurement properties to be assessed without supporting causal inference (Maxwell & Delaney, 2004). The analysis estimated the association between AE and GPA and assessed the measurement properties of the AES. It did not test causal effects, instructional mechanisms, or changes in learning over time. Gender and academic discipline were included as broad structural and contextual covariates, not as complete measures of social structure, institutional stratification, or instructional context.

The dataset comprised records from several Colombian universities and covered students from different academic disciplines and regions. This variation supported the psychometric and association analyses while allowing the models to account for gender and academic discipline. The cross-sectional structure limited the findings to evidence of association rather than evidence of causal or mediated instructional effects.

3.2. Participants

After the data were screened, the analytic sample included 1713 final-year undergraduate students. Data screening followed two

criteria: (1) completion time, in which cases with unusually short response times relative to the average survey duration were excluded; and (2) response patterns, in which cases with invariant or highly repetitive responses across items were removed.

Participants were recruited from six accredited private universities across five Colombian regions (Caribbean, Central, Western, Eastern, and South Pacific). Institutional identities were not disclosed to preserve confidentiality agreements. Recruitment occurred through institutionally coordinated invitations, including program-level communications, institutional media, student representatives, and email reminders. Participation was voluntary, and the students completed the survey online. Although the generalizability of the resulting nonprobabilistic sample was limited, it reflected the voluntary nature of participation within an institutional recruitment framework. The sample represented students from the six participating accredited private universities and was not intended to represent the Colombian higher education system as a whole.

The sample included variation across seven academic disciplines and reflected heterogeneity in regional and institutional contexts within Colombian higher education. The gender distribution included 55.46% women, 43.14% men, and 1.40% nonbinary or other identities. The mean age was 24.10 years ($SD = 5.90$). The nonbinary/other group was reported descriptively but was not included in gender-based regression or measurement invariance analyses because of its small size. The demographic characteristics of the participants are shown in Table 1.

3.3. Instruments and measures

3.3.1. Agentic engagement scale (AES)

The five-item AES developed by Reeve (2013) measures students' AE as self-reported contributions to classroom instruction rather than as student agency in its broader theoretical sense. The scale captures observable instructional behaviors through which students attempt to influence classroom instruction by expressing preferences, asking questions, and communicating needs to the instructor. These items include "I let my teacher know what I need and want" (AES1), "During this class, I express my preferences and opinions" (AES2), "When I need something in this class, I will ask the teacher for it" (AES3), "During class, I ask questions to help me learn" (AES4), and "I let my teacher know what I am interested in" (AES5).

Prior studies support the AES as a reliable and structurally coherent measure of AE. Reeve's original validation with South Korean samples revealed internal consistency between $\alpha = .77$ and $\alpha = .86$; confirmed AE as distinct from behavioral, emotional, and cognitive engagement; and showed predictive validity for perceived autonomy and achievement. In Peru, Matos et al. (2018) reported $\alpha = .85$–$.86$, temporal invariance, and validity evidence in a university sample. In Spain, Guerrero et al. (2023) reported $\alpha = .84$ and $\omega = .86$, a unidimensional structure, and validity evidence in university students. These findings supported the use of the AES, but the Colombian application still required local linguistic and cultural adaptation.

With respect to the Colombian adaptation, the original English version was used as the source version, and prior Spanish-language formulations were considered regional references. The items were translated, reviewed by bilingual experts, compared item by item, and adjusted to Colombian academic Spanish. The review focused on semantic clarity, natural wording, conceptual equivalence, and the contextual plausibility of the student–teacher behaviors described in the items. A pilot review with 20 undergraduate students

Table 1
Demographic characteristics of the participants (N = 1713).

Variable	Categories	*n*	%
Region	Caribbean	633	36.95
	Central	523	30.53
	Western	277	16.17
	Eastern	125	7.30
	South Pacific	155	9.05
Age (years)	17–20	230	13.43
	21–24	1230	71.80
	25–38	253	14.77
	$M = 24.10$, $SD = 5.90$	—	—
Academic Discipline	Arts	179	10.45
	Business & Economics	222	12.96
	Education	93	5.43
	Engineering & Architecture	326	19.03
	Health Sciences	401	23.41
	Natural Sciences & Mathematics	96	5.60
	Social Sciences & Humanities	396	23.12
GPA Range	2.1–2.5	1	0.06
	2.6–3.0	9	0.52
	3.1–3.5	55	3.21
	3.6–4.0	590	34.44
	4.1–4.5	871	50.85
	4.6–5.0	187	10.92
Gender Identity	Women	950	55.46
	Men	739	43.14
	Nonbinary/Other	24	1.40

Note. The GPA reflects Colombia's 0–5 grading scale.

examined clarity, comprehension, and contextual adequacy. Their feedback informed final wording adjustments before survey administration.

Responses were recorded on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree), where higher scores indicated higher levels of AE. The psychometric component of the study focused on the AES. The broader survey also collected data on gender, age, academic discipline, region, and self-reported cumulative GPA.

3.3.2. Academic performance

Self-reported cumulative GPA on Colombia's standardized 0–5 grading scale was used as the empirical indicator of academic performance. Because institutional records were not accessible in this multi-institutional design, self-reported GPA served as a feasible measure. Prior evidence supports this decision, showing strong correspondence between self-reported and official GPA in higher education populations ($r > .90$; Kuncel et al., 2005). Participants received instructions to report their cumulative GPA accurately during anonymous data collection. Consistent with the conceptual framing of the study, GPA was interpreted as an indicator of institutionally assessed performance, not as direct evidence of learning. The interpretation of the results accounted for potential recall or reporting effects associated with self-reported GPA.

3.4. Procedure

Data were collected electronically through an online survey administered via Microsoft Forms between March and May 2025. The participating universities distributed the survey link through program-level communications, institutional media, student representatives, and email reminders. Participation was voluntary. Before accessing the instrument, participants reviewed the study information and provided informed consent. The participants completed the questionnaire in a single session.

The survey required responses to all the items before submission. This design prevented item-level missing data but may have introduced forced-response bias. Anonymous data collection and voluntary participation were used to reduce potential response distortions in AE responses and self-reported GPA. No identifiable information was collected.

3.5. Data analysis

The analyses were conducted in R (v4.4.3; R Core Team, 2025) using RStudio and the lavaan, semTools, boot, and MASS packages. Selected R code for the main analyses is available in Section III of the Supplementary Material.

The analytical strategy addressed two objectives: (1) assessing the psychometric properties of the AES, including reliability, factorial structure, and measurement invariance, and (2) estimating the association between AE and academic performance (GPA), controlling for gender and academic discipline.

Confirmatory factor analysis (CFA) was conducted with the weighted least squares mean- and variance-adjusted (WLSMV) estimator because the AES items were ordinal Likert-type responses. WLSMV estimates factor models using polychoric correlations and does not require multivariate normality for ordinal data (Brown, 2015; Li, 2016).

For the first objective, ordinal Cronbach's α and McDonald's ω were used to estimate internal consistency, and CFA based on polychoric correlations was used to test the unidimensional structure of the AES. Model fit was evaluated using CFI ($\geq .95$), TLI ($\geq .90$), RMSEA ($\leq .08$), and SRMR ($\leq .08$). Composite reliability (CR) and average variance extracted (AVE), calculated from standardized factor loadings, were used to assess convergent validity. Multigroup CFA tested measurement invariance across women and men through configural, metric, and scalar models. The nonbinary/other group was not included in invariance testing because its small size did not support stable group-based estimation. Invariance was supported when $\Delta CFI \leq .010$ (Cheung & Rensvold, 2002).

For the second objective, multiple linear regression using ordinary least squares (OLS) was used to estimate the association between AES scores and self-reported GPA, controlling for gender and academic discipline. OLS was appropriate because GPA was treated as a continuous academic performance indicator on Colombia's 0–5 grading scale. Gender and academic discipline were entered as covariates, with academic discipline represented through dummy variables. CFA was used to assess the measurement properties of the AES, whereas regression analysis was used to estimate how AE was related to variation in academic performance. Item-level regressions were also estimated to examine whether the association with GPA reflected the overall AES score or specific agentic behaviors.

Bootstrapping with 999 resamples generated confidence intervals to reduce reliance on parametric assumptions (Hayes, 2022). Robust regression (rlm) was used to evaluate coefficient stability under nonnormality (Shapiro–Wilk $p < .001$) and heteroscedasticity (Breusch–Pagan $p < .01$). Multicollinearity diagnostics indicated no concerns (all VIFs < 2.0). OLS estimates were interpreted as the main results, and robust regression estimates and bootstrapped confidence intervals were used as sensitivity checks.

3.6. Ethical considerations

The study was approved by the Ethics Committee of the Faculty of Education at Universidad de La Sabana (Protocol No 40–2024). The procedures involving human participants followed institutional guidelines, applicable national regulations, and the ethical principles of the Declaration of Helsinki.

Participation was voluntary, and the participants provided informed consent before completing the survey. The instructions explained the study purpose, the anonymous nature of the data collection, and the participants' right to withdraw without consequences.

Data were stored anonymously, and no identifying information was linked to individual responses. The data were used exclusively for academic research purposes.

4. Results

The analysis first evaluated the psychometric properties of the AES—reliability, factorial structure, and measurement invariance across women and men—before the association between AE and academic performance (GPA) was estimated.

4.1. Measurement properties of the agentic engagement scale (AES)

4.1.1. Reliability and internal consistency

The AES exhibited high internal consistency (Table 2). The estimates from the polychoric correlation matrix converged across indices (α = .920, ω = .938), exceeding the recommended .70 threshold (Kline, 2016; Nunnally, 1967). Composite reliability (CR = .938) and average variance extracted (AVE = .751) surpassed established benchmarks, indicating that the latent construct accounted for a substantial proportion of the variance in its indicators (Watkins, 2018).

4.1.2. Factorial structure and convergent validity

CFA via the WLSMV estimator supported the unidimensional specification of the AES. Table 3 reports the fit indices, standardized factor loadings, and AVE for the model. The model yielded excellent incremental fit indices (CFI = .998, TLI = .997) and a low standardized root mean square residual (SRMR = .031) (Hu & Bentler, 1999). Given the small degrees of freedom ($df = 5$), the elevated RMSEA (.118, 90% CI [.101, .137]) was interpreted alongside the CFI, TLI, SRMR, factor loadings, reliability indices, and AVE (Kenny et al., 2015; Xia & Yang, 2019). In support of the scale's convergent validity, all standardized factor loadings were statistically significant (.795–.909), and the polychoric item correlations ranged from .635 to .809.

4.1.3. Measurement invariance across women and men

Multigroup CFA tested measurement invariance across women and men by sequentially estimating configural, metric, and scalar models. Table 4 reports the fit indices and model comparisons for each invariance step. Model comparisons revealed no degradation in fit across increasingly constrained specifications (ΔCFI = .000), satisfying the criterion for invariance (Cheung & Rensvold, 2002).

Reliability estimates remained comparable between women (α = .927, ω = .945) and men (α = .914, ω = .931). Section I of the Supplementary Material provides comprehensive psychometric results, including item-level statistics, polychoric correlations, standardized loadings, and threshold parameters (Tables S1–S5).

4.2. Association between agentic engagement and academic performance

The association between AE and academic performance was examined using OLS. The AES total score was positively associated with self-reported GPA ($N = 1689$ after excluding the nonbinary/other group from gender-based inferential analyses), with $B = 0.045$ ($SE = 0.011$, $p < .001$) and a 95% bootstrap confidence interval of [0.023, 0.068] (Table 5). The coefficient indicates a positive but modest association between overall AE and GPA.

4.2.1. Item-level analysis

Item-level disaggregation revealed that only AES2 ("During this class, I express my preferences and opinions") was associated with GPA ($B = 0.061$, $p = .001$, 95% CI [0.028, 0.094]). The remaining items (AES1, AES3, AES4, and AES5) showed no statistically significant associations. This pattern indicates that the association between AE and GPA was concentrated in students' expression of preferences and opinions rather than distributed across all AES indicators. This item-level pattern does not contradict the unidimensional psychometric structure of the AES, but it limits the interpretation of the total score as showing a uniform association with GPA.

Table 2
Reliability and convergent validity estimates for the agentic engagement scale (AES).

Statistic	Value
Number of cases (*N*)	1713
Number of items (*k*)	5
Cronbach's α (ordinal)	.920
McDonald's ω (total)	.938
Composite Reliability (CR)	.938
Average Variance Extracted (AVE)	.751

Note. All estimates are derived from the unidimensional CFA via the WLSMV estimator on polychoric correlations.

Table 3
Goodness-of-fit and convergent validity indices for the unidimensional CFA model (WLSMV estimator).

Fit index	Value obtained	Recommended value	Interpretation
χ^2 (*df*)	125.063 (5)	—	Significant; expected with large *N*
CFI (Robust)	.998	≥ .95 (Excellent)	Excellent
TLI (Robust)	.997	≥ .90 (Excellent)	Excellent
RMSEA (Robust) [90% CI]	.118 [.101, .137]	≤ .08 (Acceptable)	Elevated; model is parsimonious
SRMR	.031	≤ .08 (Good)	Excellent
Standardized Factor Loadings (λ)	.795–.909	> .70	All items strong indicators
Explained Variance (AVE)	.751	≥ .50 (Adequate)	Excellent

Note. The unidimensional model was tested via CFA.

Table 4
Measurement invariance testing of the AES across women and men.

Model	*df*	CFI	TLI	RMSEA [90% CI]	SRMR	ΔCFI vs. previous	DIFFTEST ($\Delta\chi^2$)	Invariance conclusion
Configural	10	.998	.997	.117 [.099, .136]	.031	—	—	Baseline model established
Metric	14	.998	.998	.098 [.083, .114]	.032	.000	4.041(4), $p = .400$	Full metric invariance achieved
Scalar	34	.998	.999	.062 [.052, .073]	.031	.000	16.080(20), $p = .712$	Full scalar invariance achieved

Note. Full scalar invariance was evaluated using ΔCFI ≤ .010 as the main criterion (Cheung & Rensvold, 2002). DIFFTEST results are reported as additional model-comparison information for the WLSMV models.

Table 5
Main regression model estimating the association between the AES total score and academic performance.

Variable	*B*	*SE*	*p*	95% CI Lower	95% CI Upper
Intercept	5.191	0.063	<.001	5.061	5.322
AES total score	0.045	0.011	<.001	0.023	0.068
Gender (Ref: Women)					
Men	−0.089	0.036	.013	−0.161	−0.015
Academic discipline (Ref: Health Sciences)					
Arts	0.696	0.061	<.001	0.596	0.798
Business & Economics	0.231	0.056	<.001	0.119	0.345
Education	0.710	0.079	<.001	0.569	0.853
Engineering & Architecture	0.111	0.053	.037	0.007	0.218
Natural Sciences & Mathematics	0.101	0.077	.192	−0.083	0.278
Social Sciences & Humanities	0.628	0.048	<.001	0.536	0.722

Note. Model estimated using OLS. Confidence intervals were obtained through bootstrap validation with 999 replicates. Adjusted $R^2 = .164$, $F(8, 1680) = 42.34$, $p < .001$.

4.2.2. Control variables

The GPA differed systematically across gender and academic disciplines. Men reported lower GPAs than women did ($B = -0.089$, $p = .013$, 95% CI [−0.161, −0.015]). Relative to those for Health Sciences, the GPA coefficients were greatest for Arts ($B = 0.696$, $p < .001$) and Education ($B = 0.710$, $p < .001$), whereas the other fields displayed smaller or nonsignificant differences. These differences should be interpreted as adjusted associations rather than evidence of specific disciplinary grading practices, instructional effects, or selection mechanisms.

4.2.3. Model diagnostics

Diagnostic tests detected heteroscedasticity (Breusch–Pagan $p < .01$). To address this assumption violation, the analysis employed bootstrap confidence intervals (999 resamples). The variance inflation factors remained below 2.0, indicating the absence of multicollinearity. The model accounted for 16.4% of the variance in GPA (adjusted $R^2 = .164$). Robust regression estimates were used as sensitivity checks and are reported in Section II of the Supplementary Material, together with all regression outputs and diagnostic tests (Tables S6–S7).

5. Discussion

This study examined whether AE can be measured consistently and how it relates to academic performance. The psychometric results supported the reliability and factorial structure of the AES, in line with prior work using the AES (Reeve, 2013). High internal consistency, a unidimensional structure, and scalar invariance across women and men indicate that the AES functioned comparably across these groups (Cheung & Rensvold, 2002; van de Vijver & Tanzer, 2004). This reduces the likelihood that observed gender differences are attributable to measurement nonequivalence. These results extend prior AES evidence to Colombian higher education, an underrepresented context in engagement research.

The findings refine how AE should be interpreted within the engagement–performance relationship. They support a more specific reading of AE than as an individual disposition with a uniform association with GPA. They do not address agency in its full theoretical scope. The evidence concerns AE as measured by the AES, namely, reportable contributions to instruction whose academic relevance depends on whether student initiative becomes visible, recognized, and connected to assessed performance (Reeve, 2013; Reeve et al., 2020). This interpretation is consistent with recent accounts of agency as situated and institutionally conditioned (Klemenčič, 2023; Mameli et al., 2023; Stenalt & Lassesen, 2022) while keeping the empirical claim at the level of AE.

AE is better read as an instructional expression of agency than as a direct measure of agency itself. Students may ask questions, express preferences, or communicate needs, but the academic relevance of these behaviors depends on the instructional setting in which they become available for teacher response. This interpretation also fits Kahu and Nelson's (2018) educational interface, where student action and institutional conditions meet through teaching and support. The cross-sectional design cannot identify whether instructional contexts shape AE, whether students select into settings where AE is more viable, or whether both processes operate together.

The AES total score showed a positive but modest association with GPA ($B = 0.045$), which is consistent with evidence that the engagement–performance relationship varies by dimension, measure, and context (Carini et al., 2006; Leach, 2016; Wong et al., 2024; Zhoc et al., 2019). These findings suggest that AE should not be interpreted only through its total score. Although the total AES score was positively associated with GPA, the item-level analysis revealed that this association was concentrated in students' expression of preferences and opinions. AE may be psychometrically coherent while still showing differentiated links with GPA at the behavioral level.

The item-level pattern provides the most direct evidence for a behavior-specific interpretation of AE. Only AES2, which captures students' expression of preferences and opinions, was associated with GPA. Expressing preferences and opinions can function as an instructional signal through which students make needs, priorities, or preferred ways of working visible to teachers (Reeve, 2013; Reeve et al., 2020). Because this behavior is public, interpretable, and directed toward the instructional exchange, it can provide information for feedback, clarification, or adjustments in explanations, tasks, pacing, or examples. AES2 connects student initiative with teacher responsiveness within the educational interface (Kahu & Nelson, 2018).

The nonsignificant associations for AES1, AES3, AES4, and AES5 indicate that not all agentic behaviors have the same relationship with assessed performance. Asking for support, asking questions, communicating needs, or expressing interests may still matter for learning and participation, but they did not show a direct association with GPA in this model. AES2 may be more closely tied to performance because it makes instructional preferences and opinions available for teacher recognition and response. Research on reciprocal relations between AE and autonomy support is consistent with this interpretation, because student input can elicit teacher feedback or instructional adjustment when teachers recognize it as constructive (Jang et al., 2024). Teacher beliefs and classroom conditions also matter, as student voice may be supported, ignored, or constrained depending on how teachers interpret agentic contributions (Zambrano et al., 2025).

The selective association of AES2 with GPA refines rather than rejects social–cognitive and self-regulated learning accounts. These perspectives describe agency and self-regulation as processes through which students monitor, adjust, and influence their learning conditions (Bandura, 2001, 2006; Pintrich, 2000; Zimmerman, 2000). The present findings suggest that AE becomes associated with academic performance when regulatory intentions, preferences, or needs are expressed in ways that can be recognized and acted upon within instruction. This interpretation does not imply that internal self-regulation lacks academic value. It indicates that, in this model, the GPA association appeared when student regulation entered the instructional exchange as a visible contribution.

GPA differed by gender and academic discipline, which indicates that assessed performance was not independent of students' characteristics or disciplinary location. This pattern is consistent with the view that academic performance reflects student characteristics, educational experiences, disciplinary expectations, and institutional conditions (Leach, 2016; Pascarella & Terenzini, 2005). These differences should not be read as evidence of a single mechanism. The study cannot determine whether they reflect instructional practices, grading norms, prior preparation, student self-selection, or disciplinary cultures. The results instead locate the AE–GPA relationship within student, academic, and disciplinary conditions that shape how student action is recognized and translated into assessed performance.

Cross-cultural comparisons of AE require evidence of measurement equivalence because behaviors such as expressing preferences, asking for support, or communicating needs may carry different meanings across instructional cultures. The present study provides evidence of AES invariance across women and men within one national context, but it does not establish cross-cultural invariance. The invariance evidence reported here supports gender-based comparisons within the Colombian sample, not comparisons across educational systems. Such caution is especially relevant for engagement constructs that travel across educational systems, where student voice, teacher responsiveness, and instructional authority may be organized differently (Kahu & Nelson, 2018).

5.1. Implications for policy and practice

5.1.1. Institutional assessment and policy

The association between students' expression of preferences and opinions (AES2) and GPA adds specificity to the relationship between engagement and academic outcomes. Prior work has often modeled engagement as a general predictor of outcomes (Kuh et al., 2008; Trowler, 2010), but the present results point to a more differentiated pattern. The AES total score was positively associated with GPA, whereas the item-level analysis identified a specific association for students' expression of preferences and opinions. For policy and practice, this distinction matters because aggregate engagement scores can obscure the specific conditions through which student action becomes instructionally relevant.

For institutional assessment, the association between AE and GPA ($B = 0.045$) reinforces the need to avoid interpreting academic performance as a direct expression of engagement. GPA and engagement provide useful evidence, but neither should be treated as a direct proxy for learning. Academic outcomes reflect multiple interacting influences, including student characteristics, instructional conditions, assessment practices, and institutional environments (Kuh et al., 2008; York et al., 2015). Institutional assessment should combine outcome indicators with evidence about the processes and conditions surrounding them.

Drawing on Kahu and Nelson's (2018) educational interface, institutional surveys should avoid measuring only global engagement levels. Accreditation and quality assurance processes should also examine whether institutions create conditions that enable students' expression of preferences and opinions, including student voice, feedback loops, teacher responsiveness, and institutional support for participation (Cook-Sather & Matthews, 2023; Kahu & Nelson, 2018). Student input is more useful for quality improvement when it is gathered during the course, connected to concrete learning activities, and returned to the instructional setting. Institutional evaluation should examine both reported engagement and the conditions that allow student contributions to enter instructional decision-making.

For institutions, the practical implication is to support feedback loops that make student voice actionable. These loops require low-burden mechanisms for gathering student input, faculty development on how to interpret that input, and time for instructors to respond within the course (Carless & Boud, 2018; Evans & Zhu, 2023; Steen-Utheim & Wittek, 2017). Without mechanisms for instructional response, collecting student input does not necessarily mean that it informs feedback, clarification, pacing, task design, or support. Institutional policy should protect the professional judgment of instructors while also creating expectations that student preferences and opinions are considered as evidence about learning conditions.

5.1.2. Teaching practice

In practice, the association between students' expression of preferences and opinions and GPA suggests a concrete instructional focus: creating structured opportunities for students to express preferences and opinions about course content, tasks, feedback, examples, and support. This focus is relevant to learning processes because such expression can make visible how students understand tasks, monitor progress, recognize difficulty, interpret feedback, and identify support needs (Carless & Boud, 2018; Evans & Zhu, 2023). It also links AE with agency, but in a bounded sense: AES2 represents a situated instructional expression of agency, not agency in its full theoretical scope (Reeve & Shin, 2020). Student input should guide clarification, feedback, examples, pacing, task design, and support rather than transfer course decisions to students.

Instructors can use short mid-course prompts, brief check-ins, exit tickets, anonymous questions, requests for alternative examples, or structured moments for students to identify confusing content, excessive workload, unclear feedback, or support needs. These formats matter because students do not express preferences and opinions through the same channels. Oral participation may be more accessible to students who are comfortable speaking publicly, whereas written, individual, group-based, or anonymous options can widen opportunities for students to express preferences and opinions (Cook-Sather & Matthews, 2023). The instructional value of AES2 lies in making student thinking and experience available for teacher judgment while the course is still unfolding.

Visibility is central to this strategy. Students should be able to see how their input is considered in instruction. Instructors can make responses visible by summarizing student input, explaining which suggestions can be incorporated, clarifying why some adjustments are not feasible, and showing how student preferences influence examples, activities, or feedback. Visible response closes the loop between student expression and teacher action. It also helps students understand that feedback is a dialogic process in which their questions, preferences, and opinions can shape clarification and support (Carless & Boud, 2018; Steen-Utheim & Wittek, 2017). Instructional development programs should also prepare faculty to recognize agentic signals. Student preferences, opinions, and suggestions may be missed or interpreted as disruptions when teachers are not prepared to read them as constructive instructional contributions (Zambrano et al., 2025).

These practices should remain bounded by learning outcomes, assessment criteria, academic standards, and curricular coherence. Student preferences and opinions should inform professional judgment, not replace it. A request for slower pacing, more examples, clearer feedback, or alternative forms of support may signal a remediable barrier to learning. A preference that conflicts with course goals or assessment standards may require explanation rather than adoption. The teaching implication of AES2 is that instructors should interpret student preferences and opinions as information about instructional conditions and respond in ways that preserve academic quality while making learning conditions more transparent and responsive.

5.2. Limitations and future research

The findings should be interpreted within methodological, design, and contextual constraints. First, the cross-sectional design limits causal inference about the relationship between AE and academic performance. Although the results indicate an association between students' expression of preferences and opinions (AES2) and GPA, the analysis cannot determine whether expressing preferences and opinions contributes to higher GPA, whether higher-performing students are more likely to express preferences and opinions, or whether both processes operate together. This constraint matters because educational associations may reflect reciprocal and context-dependent processes rather than unidirectional effects.

Second, reliance on self-reported measures introduces potential bias. Social desirability and recall effects may influence students' reports of AE and cumulative GPA. Although prior research has demonstrated strong correspondence between self-reported and official GPAs in higher education (Kuncel et al., 2005), measurement error cannot be fully excluded. GPA was also treated as an institutional indicator of academic performance, not as a direct measure of learning. The findings should be interpreted as evidence about assessed performance, not about learning quality, depth, or processes.

Third, sample characteristics restrict the generalizability of the findings. The research focused on students from accredited private

universities, representing a specific segment of the Colombian higher education system. Because these institutions differ from public universities in terms of resources, organizational structures, and instructional practices, such systemic differences may shape both the expression of AE and its relationship with academic performance. The study also did not include direct measures of social class, income, cultural capital, first-generation status, or prior educational trajectories. These omissions limit the analysis of structural conditions that may shape both GPA and opportunities for students to contribute to instruction. Additionally, the lower representation of disciplines such as Natural Sciences and Mathematics warrants caution when these results are extended to science, technology, engineering, and mathematics (STEM) fields, where instructional formats and participation dynamics may differ.

Fourth, additional methodological considerations concern the measurement results. The elevated RMSEA in the CFA requires cautious interpretation given its sensitivity to models with few degrees of freedom and large samples (Kenny et al., 2015). Similarly, the absence of reverse-coded items in the AES increases susceptibility to acquiescence bias (Reeve, 2013). Measurement invariance was tested across women and men but not across academic disciplines, regions, institutions, or countries. The evidence supports gender-based comparisons within this sample but not broader claims about measurement equivalence across other groups or contexts. These issues do not invalidate the findings, but they indicate areas where measurement refinement may strengthen future applications of the instrument.

Future research should address these limitations by focusing on the conditions under which students' expression of preferences and opinions (AES2) is associated with academic performance. Longitudinal designs are needed to examine whether this behavior precedes later GPA, follows prior academic performance, or develops reciprocally with academic performance over time. Multiple methods that combine self-reports with behavioral indicators, such as classroom observations and instructor reports, could clarify whether AES2 captures actual behavior, perceived opportunity, or both.

Expanding the analysis to public universities, STEM fields, and diverse instructional modalities would test whether the behavior-specific association of AES2 with GPA is generalizable across contexts. Future studies should also include structural variables such as social class, income, cultural capital, first-generation status, and prior educational trajectories. These measures would allow stronger tests of how unequal educational conditions shape both GPA and opportunities for students to contribute to instruction. Further research should also examine item-level AES results, feedback literacy processes, and instructor responsiveness. Future studies should also test whether the AES operates equivalently across disciplines, regions, institutions, institutional sectors, countries, and languages before observed differences are interpreted as substantive.

6. Conclusions

This study provides evidence that AE can be measured reliably and comparably using the AES in this Colombian higher education sample. The scale showed high internal consistency, a unidimensional structure, and scalar invariance across women and men. AE was positively and modestly associated with GPA. The item-level analysis revealed that this association was concentrated in students' expression of preferences and opinions (AES2), whereas AES1, AES3, AES4, and AES5 were not statistically associated with GPA.

These findings refine the interpretation of AE in higher education. AE can be psychometrically coherent while showing differentiated links with assessed performance at the behavioral level. Its relationship with GPA depends not only on student action in general but also on specific instructional behaviors through which students make preferences and opinions visible. This study supports a context-sensitive interpretation of AE, in which students' expression of preferences and opinions becomes instructionally relevant when instructional conditions allow that expression to be recognized and used within teaching.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author used OpenAI ChatGPT to support preliminary R code drafting and language/style revision. After using this tool, the author reviewed, edited, and verified the content as needed and takes full responsibility for the content of the manuscript.

Funding

This project was funded by the Dirección General de Investigación Universidad de La Sabana, Colombia (EDUPHD-18-2024). This article is based on the research project "Learning Assurance for the Continuous Improvement of Higher Education Institutions in Colombia: An Approach Based on Student Engagement", conducted by Uriel Eduardo Torres Castro as part of his doctoral research.

CRediT authorship contribution statement

Uriel Eduardo Torres Castro: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijer.2026.103111.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

Anzelin, I., Pineda-Báez, C., Moreno Perdomo, I. M., Galli, M., & Calderón, D. (2025). Inclusion in higher education in Argentina, Colombia, and Ecuador: Conceptualizations, advancements and challenges in access, retention, and graduation. *International Journal of Educational Research, 132*, Article 102645. https://doi.org/10.1016/j.ijer.2025.102645

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1*(2), 164–180. https://doi.org/10.1111/j.1745-6916.2006.00011.x

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education, 47*(1), 1–32. https://doi.org/10.1007/s11162-005-8150-9

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cook-Sather, A., & Matthews, K. E. (2023). Practising student voice in university teaching and learning: Three anchoring principles. *Journal of University Teaching & Learning Practice, 20*(6), 02. https://doi.org/10.53761/1.20.6.2

Evans, C., & Zhu, X. (2023). The development and validation of the assessment engagement scale. *Frontiers in Psychology, 14*, Article 1136878. https://doi.org/10.3389/fpsyg.2023.1136878

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Guerrero, E., Barrios, M., Chai, C. A., Gómez-Benito, J., Aza, A., & Guilera, G. (2023). Psychometric properties of the Spanish-language version of the Agentic Engagement Scale (AES): a preliminary study. In *International Conference on Higher Education Advances* (pp. 1477–1484). Universidad Politecnica de Valencia. https://doi.org/10.4995/HEAd23.2023.16122.

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). The Guilford Press.

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

Jang, H.-R., Basarkod, G., Reeve, J., Marsh, H. W., Cheon, S. H., & Guo, J. (2024). Longitudinal reciprocal effects of agentic engagement and autonomy support: Between- and within-person perspectives. *Journal of Educational Psychology, 116*(1), 20–35. https://doi.org/10.1037/edu0000815

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development, 37*(1), 58–71. https://doi.org/10.1080/07294360.2017.1344197

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

Kingsford-Smith, A. A., Evans, P., & Jung, J. Y. (2026). Engaged or just compliant? A person-centered investigation of the dimensionality and structure of student engagement. *Educational Psychology* (pp. 1074–1093). https://doi.org/10.1080/01443410.2026.2679186

Klemenčič, M. (2023). A theory of student agency in higher education. In C. Baik, & E. R. Kahu (Eds.), *Research handbook on the student experience in higher education* (pp. 25–40). Edward Elgar Publishing. https://doi.org/10.4337/9781802204193.00010.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Koper, P. T., Felton, J., Sanney, K. J., & Mitchell, J. B. (2015). Real GPA and real SET: Two antidotes to greed, sloth and cowardice in the college classroom. *Assessment & Evaluation in Higher Education, 40*(2), 248–264. https://doi.org/10.1080/02602938.2014.904272

Kuh, G. D. (2003). What we're learning about student engagement from NSSE: Benchmarks for effective educational practices. *Change: The Magazine of Higher Learning, 35*(2), 24–32. https://doi.org/10.1080/00091380309604090

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education, 79*(5), 540–563. https://doi.org/10.1080/00221546.2008.11772116

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*(1), 63–82. https://doi.org/10.3102/00346543075001063

Labraña, J., & Brunner, J. J. (2022). Transformación de la educación superior latinoamericana y su impacto en la idea de la universidad: Del acceso de élite a la masificación y universalización del acceso [Transformation of Latin American higher education and its impact on the idea of the university: From elite access to massification and universalization of access]. *Perfiles Educativos, 44*(176), 138–151. https://doi.org/10.22201/iisue.24486167e.2022.176.60539

Leach, L. (2016). Exploring discipline differences in student engagement in one institution. *Higher Education Research & Development, 35*(4), 772–786. https://doi.org/10.1080/07294360.2015.1137875

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: An International Journal, 46*(3), 517–528. https://doi.org/10.2224/sbp.7054

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Mameli, C., Grazia, V., & Molinari, L. (2023). Student agency: Theoretical elaborations and implications for research and practice. *International Journal of Educational Research, 122*, Article 102258. https://doi.org/10.1016/j.ijer.2023.102258

Martínez, A., Borjas, M., Herrera, M., & Valencia, J. (2015). Relationship between measures of academic quality and undergraduate student attrition: The case of higher education institutions in the Colombian Caribbean region. *Higher Education Research & Development, 34*(6), 1192–1206. https://doi.org/10.1080/07294360.2015.1024622

Matos, L., Reeve, J., Herrera, D., & Claux, M. (2018). Students' agentic engagement predicts longitudinal increases in perceived autonomy-supportive teaching: The squeaky wheel gets the grease. *The Journal of Experimental Education, 86*(4), 579–596. https://doi.org/10.1080/00220973.2018.1448746

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Lawrence Erlbaum Associates Publishers.

Melguizo, T., Sanchez, F., & Velasco, T. (2016). Credit for low-income students and access to and academic performance in higher education in Colombia: A regression discontinuity approach. *World Development, 80*, 61–77. https://doi.org/10.1016/j.worlddev.2015.11.018

Michou, A., Altan, S., Mouratidis, A., Reeve, J., & Malmberg, L. E. (2023). Week-to-week interplay between teachers' motivating style and students' engagement. *The Journal of Experimental Education, 91*(1), 166–185. https://doi.org/10.1080/00220973.2021.1897774

Moreira, P. A. S., Cunha, D., & Inman, R. A. (2020). An integration of multiple student engagement dimensions into a single measure and validity-based studies. *Journal of Psychoeducational Assessment, 38*(5), 564–580. https://doi.org/10.1177/0734282919870973

Nunnally, J. C. (1967). *Psychometric theory*. McGraw-Hill.

Ogunsakin, R. E., Moyo, S., Olugbara, O. O., & Israel, C. (2021). Relating student engagement indicators to academic performance using multiple correspondence analysis. *Cybernetics and Information Technologies, 21*(1), 87–102. https://doi.org/10.2478/cait-2021-0007

Páez, P. N., & Teelken, C. (2016). Perceived quality of private education and fears of stratification: Investigating the propositions of human capital theory by exploring the case of Colombia. *Policy Futures in Education, 14*(6), 597–616. https://doi.org/10.1177/1478210316639416

Parra Sandoval, M. C. (2021). Perfil del estudiante universitario latinoamericano [Profile of the Latin American university student]. Oficina para América Latina del Instituto. *Internacional de Planeamiento de la Educación de la UNESCO*. https://unesdoc.unesco.org/ark:/48223/pf0000378123.

Pascarella, E. T., & Terenzini, P. T. (2005). How college affects students: A third decade of research. *Jossey-Bass*.

Patall, E. A., Kennedy, A. A. U., Yates, N., Zambrano, J., Lee, D., & Vite, A. (2022). The relations between urban high school science students' agentic mindset, agentic engagement, and perceived teacher autonomy support and control. *Contemporary Educational Psychology, 71*, Article 102097. https://doi.org/10.1016/j.cedpsych.2022.102097

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. The National Academies Press. https://doi.org/10.17226/10019.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3.

R Core Team. (2025). *R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing*. https://www.R-project.org/.

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology, 105*(3), 579–595. https://doi.org/10.1037/a0032690

Reeve, J., Cheon, S. H., & Jang, H. (2020). How and why students make academic progress: Reconceptualizing the student engagement construct to increase its explanatory power. *Contemporary Educational Psychology, 62*, Article 101899. https://doi.org/10.1016/j.cedpsych.2020.101899

Reeve, J., & Jang, H. (2022). Agentic engagement. In A. L. Reschly, & S. L. Christenson (Eds.), *Handbook of research on student engagement* (2nd ed., pp. 95–107). Springer. https://doi.org/10.1007/978-3-031-07853-8_5.

Reeve, J., Jang, H., Shin, S. H., Ahn, J. S., Matos, L., & Gargurevich, R. (2022). When students show some initiative: Two experiments on the benefits of greater agentic engagement. *Learning and Instruction, 80*, Article 101564. https://doi.org/10.1016/j.learninstruc.2021.101564

Reeve, J., & Shin, S. H. (2020). How teachers can support students' agentic engagement. *Theory Into Practice, 59*(2), 150–161. https://doi.org/10.1080/00405841.2019.1702451

Reeve, J., & Tseng, C. M. (2011). Agency as a fourth aspect of student engagement during learning activities. *Contemporary Educational Psychology, 36*(4), 257–267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387. https://doi.org/10.1037/a0026838

Rodríguez-Hernández, C. F., Cascallar, E., & Kyndt, E. (2020). Socio-economic status and academic performance in higher education: A systematic review. *Educational Research Review, 29*, Article 100305. https://doi.org/10.1016/j.edurev.2019.100305

Steen-Utheim, A., & Wittek, A. L. (2017). Dialogic feedback and potentialities for student learning. *Learning, Culture and Social Interaction, 15*, 18–30. https://doi.org/10.1016/j.lcsi.2017.06.002

Steenman, S. C., Bakker, W. E., & van Tartwijk, J. W. F. (2016). Predicting different grades in different ways for selective admission: Disentangling the first-year grade point average. *Studies in Higher Education, 41*(8), 1408–1423. https://doi.org/10.1080/03075079.2014.970631

Stenalt, M. H., & Lassesen, B. (2022). Does student agency benefit student learning? A systematic review of higher education research. *Assessment & Evaluation in Higher Education, 47*(5), 653–669. https://doi.org/10.1080/02602938.2021.1967874

Tomkin, J. H., & West, M. (2022). STEM courses are harder: Evaluating inter-course grading disparities with a calibrated GPA model. *International Journal of STEM Education, 9*, 27. https://doi.org/10.1186/s40594-022-00343-1

Torres Castro, U. E., & Pineda-Baéz, C. (2026). Student agency in Colombian higher education: A dual-pathway model of mediation and moderation between student engagement and academic achievement. *Higher Education*. https://doi.org/10.1007/s10734-026-01669-3

Trowler, V. (2010). *Student engagement literature review. The Higher Education Academy*. https://www.improvingthestudentexperience.com/wp-content/uploads/library/UG_documents/StudentEngagementLiteratureReview_2010.pdf.

van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology, 54*(2), 119–135. https://doi.org/10.1016/j.erap.2003.12.004

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology, 44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Westrick, P. A. (2017). Reliability estimates for undergraduate grade point average. *Educational Assessment, 22*(4), 231–252. https://doi.org/10.1080/10627197.2017.1381554

Wong, Z. Y., Liem, G. A. D., Chan, M., & Datu, J. A. D. (2024). Student engagement and its association with academic achievement and subjective well-being: A systematic review and meta-analysis. *Journal of Educational Psychology, 116*(1), 48–75. https://doi.org/10.1037/edu0000833

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods, 51*(1), 409–428. https://doi.org/10.3758/s13428-018-1055-2

York, T. T., Gibson, C., & Rankin, S. (2015). Defining and measuring academic success. *Practical Assessment, Research, and Evaluation, 20*(1), 5. https://doi.org/10.7275/hz5x-tx03

Zambrano, J., Patall, E. A., Kennedy, A. A. U., Aguilera, C., & Yates, N. (2025). Qualitative study of urban high school teachers' beliefs about students' agentic engagement. *The Journal of Experimental Education, 93*(1), 69–90. https://doi.org/10.1080/00220973.2023.2238632

Zhoc, K. C. H., Webster, B. J., & King, R. B. (2019). Higher education student engagement scale (HESES): Development and psychometric evidence. *Research in Higher Education, 60*, 219–244. https://doi.org/10.1007/s11162-018-9510-6

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7.