MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2026, 22(9), em2917
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/19318

Validation and adaptation of the epistemic beliefs survey for mathematics in Colombian secondary students

Eder Santiago Martelo-Gomez [1*], Eduin Alexander Prada-Rodríguez [1], Rosa Delgado-Rebolledo [2]

[1] Universidad de Investigación y Desarrollo, Bucaramanga, COLOMBIA
[2] Universidad de Valparaíso, Valparaíso, CHILE

Received 20 April 2026 ▪ Accepted 08 August 2026

Abstract

This study adapted and validated the epistemic beliefs survey for mathematics for Colombian secondary students. A total of 817 public-school students aged 10-20 participated, and the sample was randomly divided for exploratory factor analysis (EFA) (n = 407) and confirmatory factor analysis (CFA) (n = 410). Content validity was examined through expert judgment, and analyses used procedures appropriate for ordinal data. Item screening and EFA supported a revised 27-item structure in which ability differentiated into personal ability and general ability, whereas structure of knowledge did not emerge as a coherent factor. CFA showed that the revised six-factor model fit the data better than the five-factor alternative (CFI = .966, TLI = .962, RMSEA = .070, SRMR = .072), with satisfactory reliability (ωt = .898). Measurement invariance held through the scalar level across educational levels. Findings suggest context-sensitive factorial configurations, supporting context-specific validation when transferring domain-specific instruments.

Keywords: epistemic beliefs, mathematics education, scale adaptation, confirmatory factor analysis, secondary education, Latin America

INTRODUCTION

Education is widely recognized as a key driver of social development, as the formation of a skilled workforce fosters economic growth, expands learning opportunities, and promotes equity and social inclusion (World Bank, 2024). In this context, mathematics plays a central role, providing fundamental tools for problem solving, logical thinking, and abstract reasoning, which are essential for individual and social development (Boaler, 2022; Cresswell & Speelman, 2020; Jablonski, 2025).

Despite its importance, mathematical performance in Latin America and the Caribbean remains a major concern. According to the 2022 program for international student assessment, the region achieved an average of 373 points in mathematics, significantly below the Organization for Economic Co-operation and Development (OECD) average of 475 points. This gap corresponds to an estimated delay of approximately five years of schooling. Furthermore, only 25% of students reached the minimum level of mathematical proficiency, and no country in the region demonstrated an equitable education system, reflecting persistent structural inequalities (Arias Ortiz et al., 2024). In Colombia, the situation follows a similar pattern, with only 29% of students reaching Level 2 proficiency and an average score of 383 points (OECD, 2023).

While these results describe students' mathematical performance, they do not fully capture the cognitive and epistemic processes underlying learning. Research has shown that students' beliefs about knowledge and learning play a critical role in shaping how they approach mathematical tasks, interpret information, and regulate their learning strategies (Schiefer et al., 2022). In particular, epistemic beliefs influence how students justify procedures, evaluate solutions, and assign meaning to mathematical knowledge (Hofer & Pintrich, 1997; Muis, 2004; Ondaro & Dagoc, 2025). Naïve beliefs, such as conceiving mathematics as a fixed set of rules or

✉ emartelo1@udi.edu.co **(*Correspondence)** ✉ eprada1@udi.edu.co ✉ rosa.delgado@uv.cl

Contribution to the literature

- This study provides evidence for the reliability and construct validity of the epistemic beliefs survey for mathematics (EBSM) in a Latin American secondary education context, where validated instruments remain limited.
- The findings show that the structure of mathematical knowledge did not emerge as a coherent dimension in this sample, raising questions about whether this construct is equally salient across educational levels and cultural contexts.
- This study proposes a refined 27-item factorial structure that differentiates ability-related beliefs into personal and general ability, contributing to a more precise understanding of students' epistemic beliefs in mathematics.

absolute truths, are associated with surface-level strategies and lower performance, whereas more sophisticated beliefs support deeper understanding and self-regulated learning (Hofer & Pintrich, 1997; Muis, 2004).

In this context, the study of epistemic beliefs in mathematics has gained increasing attention. Early research referred to epistemological beliefs (Hofer & Pintrich, 1997; Schommer, 1990), but more recent work has adopted the term epistemic beliefs to emphasize their psychological and educational nature (Kelly, 2020; Sandoval et al., 2016). These beliefs refer to students' conceptions about the nature of mathematical knowledge and the processes through which it is constructed, justified, and validated (Hofer & Pintrich, 2002).

Research on epistemic beliefs has evolved from developmental models (Perry, 1970) to multidimensional approaches that conceptualize beliefs as distinct but related dimensions (Schommer, 1990). However, instruments derived from these frameworks have shown limitations in terms of factorial validity, structural stability, and replicability across contexts (Alkış Küçükaydin & Ayaz, 2024; Dorsah et al., 2025; Říčan & Pešout, 2022; Schraw et al., 2002; Walker, 2007). These challenges highlight the need for rigorous psychometric analyses and contextual adaptations when instruments are applied in different educational and cultural settings.

In response to the need for domain-specific measurement, Walker (2007) developed the EBSM, a multidimensional instrument that assesses six dimensions: source of knowledge, certainty of knowledge, structure of knowledge, speed of knowledge, innate ability, and real-world applicability. Although this instrument has been widely used, it was originally developed in an English-speaking cultural context, which may limit its applicability in other educational settings due to possible differences in item interpretation and factorial structure (Brown, 2015; Schraw et al., 2002).

Psychometric research indicates that the transfer of instruments across cultural contexts may affect item interpretation, factorial structure, and measurement validity (Říčan & Pešout, 2022; Schraw et al., 2002). Evidence from studies conducted in Colombia suggests that the theoretical structure of belief scales is not always replicated, requiring adjustments and reorganization of factors (Rincón-Álvarez et al., 2022). Despite this, there is still limited empirical evidence regarding the functioning of the EBSM in Latin American contexts, particularly in secondary education.

A central limitation in the current literature is the lack of robust evidence on the factorial structure and psychometric functioning of the EBSM in Latin American secondary education. Although the instrument has been used to examine students' epistemic beliefs, there is limited evidence based on advanced psychometric procedures, such as exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and systematic model comparison. This gap is particularly relevant because the EBSM was originally developed in an English-speaking cultural context, and its application in other settings requires careful cultural adaptation and validation. Psychometric research has consistently shown that transferring instruments across cultural contexts can affect item interpretation, factorial structure, and measurement validity (Říčan & Pešout, 2022; Schraw et al., 2002). Evidence from Colombia reinforces this concern: Rincón-Álvarez et al. (2022) found that the theoretical structure of a mathematics belief scale was not fully reproduced among Colombian students.

Consequently, it remains unclear whether the original factorial structure of the EBSM can be replicated, whether it requires partial adaptation, or whether it reflects context-dependent reconfigurations of students' epistemic belief systems. This limitation constrains not only the interpretation of scores obtained from the instrument, but also the theoretical understanding of whether epistemic beliefs in mathematics operate as independent dimensions or as a more flexible system of interrelated beliefs shaped by educational and cultural experience.

In line with this rationale, this study aims to culturally adapt and validate the EBSM and to examine its underlying factorial structure in Colombian secondary students, with a focus on dimensionality, construct validity, and the extent to which its structure may reflect context-dependent configurations of epistemic beliefs.

THEORETICAL FRAMEWORK

Epistemic beliefs refer to individuals' conceptions about the nature of knowledge and how it is constructed and justified (Hofer & Pintrich, 2002; Muis, 2004). In educational contexts, these beliefs shape how students interpret tasks, evaluate information, and regulate their learning strategies. In mathematics education, they are particularly relevant because they influence how students understand reasoning, problem solving, and the legitimacy of procedures.

Students who conceive mathematics as a fixed body of rules determined by authority tend to rely on procedural and surface-level approaches to learning. In contrast, those who view mathematical knowledge as interconnected and evolving are more likely to engage in conceptual understanding and reflective reasoning (Hofer & Pintrich, 1997; Muis, 2004). Accordingly, epistemic beliefs constitute a key construct for understanding how students approach mathematical tasks and engage with mathematical knowledge. This perspective is also consistent with Ernest's (1989) distinction between instrumentalist, Platonist, and problem-solving views of mathematics, which provides a useful theoretical lens for understanding whether mathematical knowledge is conceived as a set of rules, as a fixed body of truths, or as a dynamic human activity centered on problem solving.

The systematic study of epistemic beliefs can be traced back to Perry (1970), who conceptualized epistemological development as a staged progression in their understandings of the certainty, complexity, and justification of knowledge. Although this model provided a foundational basis for subsequent research on epistemic cognition, its sequential and holistic nature posed challenges for empirical measurement, since beliefs were conceived as part of an overarching trajectory rather than as independent, measurable dimensions (Hofer & Pintrich, 1997; Muis, 2004).

Building on this foundation, Schommer (1990) proposed a multidimensional model that conceptualized epistemic beliefs as a system of relatively independent but interrelated dimensions, encompassing beliefs about both the nature of knowledge and the processes of knowing. Despite its broad theoretical influence, instruments derived from this model have shown recurring limitations in factorial stability and internal consistency (Schraw et al., 2002; Walker, 2007). Recent studies suggest that these challenges persist with problems related to factorial validity and limited replicability across educational and cultural contexts (Alkış Küçükaydin & Ayaz, 2024; Dorsah et al., 2025; Říčan & Pešout, 2022). In response to the limitations of general instruments, Walker (2007) developed a domain-specific multidimensional model of epistemic beliefs in mathematics. This model draws on earlier theoretical frameworks (Hofer & Pintrich, 1997; Schommer, 1990) and incorporates a dimension of real-world applicability informed by research on mathematical beliefs (Schoenfeld, 1989). From this model, Walker constructed the EBSM, an instrument organized around six broad dimensions: source of knowledge, certainty of knowledge, structure of knowledge, speed of knowledge acquisition, innate ability, and real-world applicability. Within the ability domain, the item content includes both beliefs about one's own mathematical ability and beliefs about mathematical ability in general, a distinction that becomes relevant for the empirical analyses reported below.

Dimensions of the Epistemic Beliefs Survey for Mathematics

The EBSM operationalizes epistemic beliefs in mathematics through six broad domains: source of knowledge, certainty of knowledge, structure of knowledge, speed of knowledge acquisition, innate ability, and real-world applicability (Walker, 2007). Source of knowledge refers to beliefs about whether mathematical knowledge is primarily transmitted by external authorities, such as teachers or textbooks, or actively constructed by learners. Certainty of knowledge captures the extent to which students view mathematical knowledge as fixed, absolute, and unchanging, as opposed to being open to justification, interpretation, and development. Structure of knowledge concerns whether mathematics is perceived as a set of isolated facts and procedures or as an interconnected system of concepts and relationships. These first three dimensions are consistent with multidimensional models of epistemic beliefs that distinguish beliefs about the source, certainty, and structure of knowledge (Hofer & Pintrich, 1997; Muis, 2004).

Speed of knowledge acquisition reflects beliefs about whether mathematical understanding should occur quickly or may require sustained effort, time, and gradual development. Innate ability refers to beliefs about whether mathematical competence is fixed and determined by natural talent or can be developed through learning and effort. This broader domain includes item content that distinguishes beliefs about one's own mathematical ability from more general beliefs about the nature and distribution of mathematical talent. Finally, real-world applicability addresses students' beliefs about the usefulness and relevance of mathematics beyond the classroom, particularly in everyday life, other school subjects, and future professional contexts (Schoenfeld, 1989; Walker, 2007).

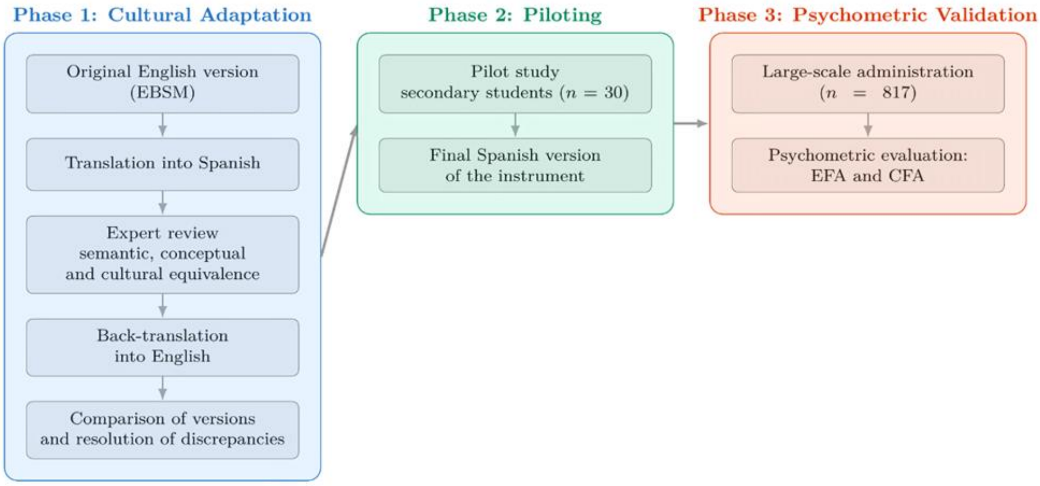

Figure 1. Flow diagram: cultural adaptation and validation process of the EBSM (the authors' own elaboration)

METHOD

Design

This study employed an instrumental, cross-sectional design to adapt the instrument into Spanish and evaluate its psychometric properties. A quantitative approach was used, combining EFA and CFA to examine item quality, internal structure, and reliability.

Participants

The sample consisted of 817 secondary education students enrolled in grades six through eleven in public schools in the metropolitan area of Bucaramanga, Colombia. Participants' ages ranged from 10 to 20 years. Mean age increased progressively by grade level, from 12.03 years in sixth grade to 16.45 years in eleventh grade. Regarding sex composition, 54.94% of the participants identified as female and 45.06% as male. Participants were selected using non-probabilistic purposive sampling. Given the applied nature of the study, purposive sampling was considered appropriate to ensure access to the target population.

Inclusion criteria were:

(a) current enrollment in the institution and participation in mathematics as part of the official curriculum, and

(b) provision of informed consent signed by parents or legal guardians.

Students with diagnoses associated with cognitive difficulties were excluded, as were those who, despite having parental consent, did not provide their own assent.

Instruments

The original English version of the EBSM (Walker, 2007) consists of 39 ordinal polytomous items with five response options ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The instrument was designed to assess six theoretical dimensions: source of knowledge (items 1-5), certainty of knowledge (items 6-9), structure of knowledge (items 10-13), speed of knowledge (items 14-20), innate ability (items 21-31), and applicability of knowledge (items 32-39). The adaptation and use of the instrument were conducted for academic research purposes, with appropriate acknowledgment of the original source. The cultural adaptation process followed the methodological guidelines described later, and the final validated structure, including the retained items in Spanish, is presented in the results section and **Appendix A**. **Appendix B** shows the supplementary tables.

Adaptation and translation into Spanish

The Spanish adaptation process followed the methodological recommendations of Hambleton and Zenisky (2011) and the guidelines of the International Test Commission (2018). Translation and back-translation were carried out by language specialists from the participating institution and subsequently reviewed by members of research groups in psychology, education and basic sciences to ensure semantic, conceptual, and cultural equivalence.

A pilot study was conducted with 30 secondary education students to examine item comprehension, clarity of instructions, and cultural relevance, contributing to the refinement of the final version of the instrument prior to its large-scale administration. The stages of the cultural adaptation and validation process are summarized in **Figure 1**.

Analytical strategy

Analyses were conducted in R (RStudio) using the psych, GPArotation, lavaan (Rosseel, 2012), and semTools packages. Preliminary analyses included descriptive statistics, item discrimination indices, and distributional indicators.

For the EFA, a polychoric correlation matrix was estimated, using unweighted least squares (ULS) extraction and oblimin rotation. Sampling adequacy was assessed with the Kaiser-Meyer-Olkin (KMO) index and Bartlett's test of sphericity ($p < .001$). Internal consistency was evaluated using ordinal alpha (a) and omega (ω) coefficients.

The CFA was conducted using the WLSMV estimator for ordinal data. Model fit was assessed using the chi-square statistic (χ^2), comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA; 90% confidence interval [CI]), and standardized root mean square residual (SRMR), following established criteria (Byrne, 2016; Hu & Bentler, 1999; Kline, 2023). In addition to absolute fit indices, competing CFA models were compared using changes in CFI and RMSEA. Δ values were calculated as the fit index of the target model minus the fit index of the reference model. Statistical significance was set at $p < .05$.

Measurement invariance across educational levels (basic secondary, grades 6-9, vs. upper secondary, grades 10-11) was examined using a sequential configural-metric-scalar-strict framework with WLSMV estimation, evaluated through scaled ΔCFI and ΔRMSEA criteria (Byrne, 2016; Kline, 2023).

Ethical Guidelines

The study was conducted in accordance with the ethical guidelines of the American Psychological Association (2017) and the principles established in the Declaration of Helsinki (World Medical Association, 2013). Ethical approval was obtained from the institutional ethics committee of the participating university.

Following institutional authorization and the completion of a pilot phase to ensure semantic and cultural adequacy of the instrument, permission was obtained from participating schools. Informed consent was collected from parents or legal guardians, and student assent was obtained prior to participation. The final version of the scale was administered in group settings within controlled academic contexts by trained survey administrators from the participating schools. Standardized training ensured consistency in instructions, timing, and handling of questions, thereby reducing common method bias and enhancing procedural reliability (Podsakoff et al., 2012). All data were handled in accordance with confidentiality principles throughout the research process.

RESULTS

To evaluate the psychometric properties of the adapted scale, EFA was conducted first to examine the underlying factor structure, followed by CFA to test competing measurement models, in accordance with the methodological criteria described in the analytical strategy.

Content Validity

Content validity of the adapted version was examined using Aiken's V index, based on the ratings provided by 11 expert judges for the 39 theoretical items across four criteria: clarity, coherence, relevance, and representativeness. An acceptance threshold of .75 was adopted. The distribution of coefficients indicated that most items obtained V values above .90 and that, on average, each criterion reached values greater than .94, suggesting a high level of agreement among judges regarding item adequacy.

Only one item yielded a V value below .75 for the clarity criterion (item 19). Its original wording was identified as overly dichotomous, suggesting an all-or-nothing interpretation of students' understanding. In response, the item was reworded to reduce the possibility of a dichotomous interpretation while preserving the underlying construct, thereby improving semantic precision and comprehensibility for students in basic and secondary education. No other items required substantive modifications; therefore, the obtained coefficients support the content validity of the adapted scale.

Data Screening

The dataset was cleaned and updated using a simple item-level mode imputation procedure for missing values observed in two items, provided that the proportion of missing data was below 10% per participant. In addition, absolute and relative frequencies were computed to identify items with limited variability; no items were found for which a single response option exceeded 95% of endorsements within a category.

The total number of valid cases was 817. These cases were randomly split into two independent subsamples using a stratified procedure by grade level: 407 cases were allocated to the EFA and 410 cases to the CFA.

Initial Exploratory Factor Analysis (39-Item Version)

An initial EFA was conducted on the full 39-item version of the instrument using a polychoric correlation matrix, ULS extraction, and oblimin rotation. This analysis aimed to examine the underlying dimensional structure prior to item refinement.

Sampling adequacy was good (KMO = .867), and Bartlett's test of sphericity was significant, supporting the factorability of the data. The model showed acceptable global fit (RMSEA = .059), although the Tucker-Lewis Index (TLI = .821) indicated that the structure did not fully meet conventional criteria for optimal fit.

Inspection of the factor loading pattern revealed several psychometric issues. While some factors exhibited strong and well-defined loadings, a substantial number of items showed weak communalities, cross-loadings, or high complexity indices, indicating ambiguous relationships with the latent dimensions. Certain items displayed low shared variance (e.g., communalities below .30), whereas others loaded simultaneously on multiple factors, compromising interpretability.

Additionally, some theoretically defined dimensions did not emerge as clearly differentiated constructs, suggesting limited internal coherence and overlap with other factors. These findings indicated that the original theoretical structure was not fully supported in this population.

Taken together, the results of this initial EFA provided empirical justification for a systematic refinement process and directly guided the identification and removal of weak items and poorly defined dimensions, with the aim of improving dimensional clarity, internal consistency, and the overall psychometric performance of the instrument prior to subsequent structural analyses.

Item Refinement

Following the initial EFA, a unidimensional analysis was conducted for each subscale to further examine the correspondence between the theoretical item allocation and the empirical structure, and to identify potentially weak items. To this end, polychoric correlations, the KMO measure of sampling adequacy, and the ordinal alpha coefficient were computed for each subscale. Bartlett's test of sphericity was significant ($p < .001$), confirming the factorability of the correlation matrices.

Preliminary analyses indicated acceptable sampling adequacy and internal consistency across most subscales. Source of knowledge (KMO = .66, $\alpha = .63$), certainty of knowledge (KMO = .66, $\alpha = .62$), innate ability (KMO = .87, $\alpha = .86$), and real-world applicability (KMO = .75, $\alpha = .67$) showed moderate to high adequacy, whereas speed of knowledge (KMO = .71, $\alpha = .23$) required further refinement.

In contrast, the structure of knowledge subscale (KMO = .49, $a = .34$) demonstrated substantially lower values, indicating limited communality and weak inter-item covariance. The mean inter-item correlation was very low ($\bar{r} = .11$), with 66.7% of correlations below .30, and a high determinant (.75), reflecting substantial internal independence. In line with established psychometric guidelines (Howard, 2016; Watkins, 2020), this dimension was removed from the model due to insufficient shared variance to support a coherent factor.

Applying these same refinement criteria (low communality, insufficient inter-item correlations, unstable loadings, or negative contributions to internal consistency), the remaining subscales were systematically examined.

Within the *source of knowledge* subscale (items 1-5), item 4 exhibited negligible associations with the remaining items (ranging from -.06 to .16), and 60% of inter-item correlations were below .30. It was therefore removed, resulting in a more coherent four-item configuration.

Within the *speed of knowledge* dimension, all possible combinations of three or more items were systematically evaluated to identify the subset that maximized internal consistency. The combination consisting of items 16, 18, and 20 yielded the highest ordinal reliability ($a = .686$) and demonstrated the most coherent inter-item associations within the dimension. Items that reduced internal homogeneity were therefore excluded.

For the *real-world applicability* subscale, items 32 and 33 showed negative loadings and extremely low communalities ($h^2 = .023$ and $h^2 = .014$), with near-zero item-total correlations. Both were removed.

Within the *innate ability* subscale, two coherent groupings emerged, corresponding to personal ability (items 22-26) and general ability (items 27-31). These results suggested a potential differentiation within the ability domain, which was further evaluated in the subsequent CFA.

The refinement process resulted in the removal of 12 items, yielding a final 27-item version. The refined version demonstrated adequate sampling adequacy (KMO = .854) and a significant Bartlett's test ($\chi^2[351] = 4{,}429.45$, $p < .001$), with satisfactory overall reliability ($\omega = .902$; $a = .702$).

Following this refinement, an EFA was conducted on the 27-item version to examine its latent structure. The overall process of psychometric refinement and structural evaluation is summarized in **Figure 2**.

Exploratory Factor Analysis (27-Item Version)

An EFA was conducted on the refined 27-item version of the instrument to examine its latent structure. Although the scale was originally grounded in a theoretical six-factor model, the EFA aimed to determine whether this structure was replicated or modified in the present context.

Sampling adequacy was very good (KMO = .854), and Bartlett's test of sphericity was significant ($\chi^2[351] = 4{,}429.45$, $p < .001$), supporting the suitability of the polychoric correlation matrix for factor analysis.

Factor retention was guided by multiple criteria, including eigenvalues, scree plot inspection, and interpretability of the factor solution. The first five eigenvalues were above 1.00 (6.824, 3.615, 1.893, 1.647, and 1.194), whereas the sixth eigenvalue was close to unity (.987), suggesting that both five- and six-factor solutions were plausible.

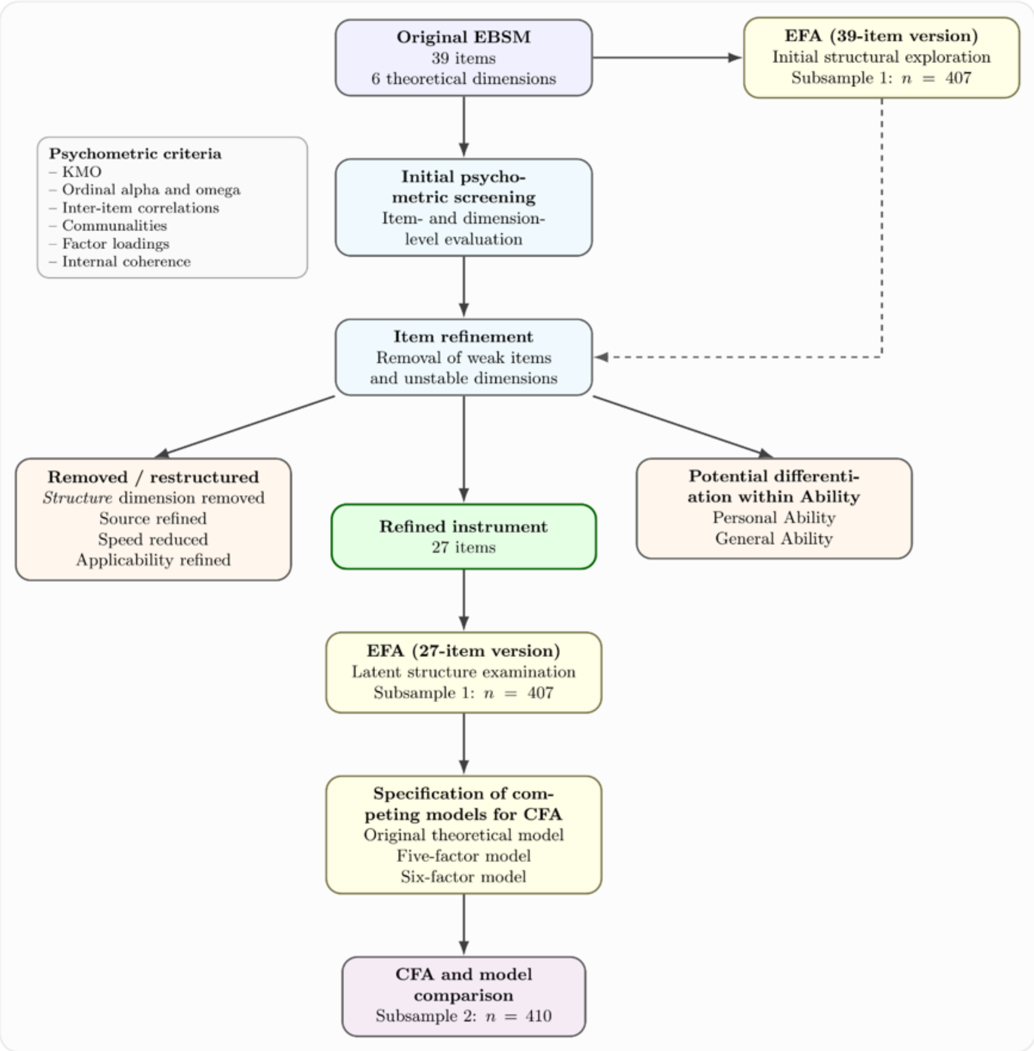

Figure 2. Flow diagram of the psychometric refinement and structural evaluation process of the EBSM (the authors' own elaboration)

The five-factor solution explained 46.6% of the variance, whereas the six-factor solution explained 48.6%. Substantively, the five-factor solution reflected the exclusion of the *structure* dimension, whereas the six-factor solution additionally differentiated the original *ability* domain into two distinct but related components: *personal ability* and *general ability*. Because the six-factor solution provided a more differentiated and interpretable representation of the data, it was retained for subsequent analyses.

Factor loadings in the retained six-factor solution were generally adequate, with most primary loadings exceeding .40. Higher loadings were observed in the ability-related dimensions, whereas comparatively weaker loadings appeared in the *certainty* and *speed* factors. Two items (CEM07 and CEM34) exhibited cross-loadings, and one item (CEM06) showed consistently low loadings across factors, indicating comparatively weaker performance.

Item communalities ranged from .24 to .70, indicating moderate to adequate shared variance across factors. Only item 6 showed a communality slightly below the conventional threshold (h^2 = .24); however, it was retained due to its conceptual relevance within the *certainty* dimension and because the item set remained theoretically interpretable (Kline, 2023).

Global reliability for the refined 27-item pool was acceptable to high, with an ordinal alpha of .702 and an omega total of .902. At the dimensional level, internal consistency was stronger for the ability-related and applicability dimensions, whereas *certainty* and especially *speed* showed comparatively weaker reliability. Interfactor correlations were generally low to moderate in magnitude, indicating that the factors were related but empirically distinguishable.

In summary, the EFA supported a six-factor structure that partially diverged from the original theoretical model. Specifically, the *structure* dimension was not retained, whereas the *ability* domain differentiated into *personal ability* and *general ability*. These findings informed the specification of three competing structural models:

Table 1. Fit indices for the theoretical and refined CFA models

Model	χ^2(df)	χ^2/df	CFI	TLI	RMSEA	SRMR
Original six-factor (39 items)	3,764.54 (687)	5.48	.622	.592	.105	.120
Five-factor (27 items)	1,148.76 (314)	3.66	.955	.950	.081	.078
Six-factor (27 items)	935.84 (309)	3.03	.966	.962	.070	.072
Criterion	-	≤ 3	≥ .95	≥ .95	≤ .08	≤ .08

Table 2. Comparative differences in fit indices across CFA models

Model comparison	Reference model	Target model	ΔCFI	ΔRMSEA	Supported model
Five-factor model, 27 items, vs. original six-factor model, 39 items	Original six-factor, 39 items	Five-factor, 27 items	+.333	-.024	Five-factor, 27-item model
Six-factor model, 27 items, vs. original six-factor model, 39 items	Original six-factor, 39 items	Six-factor, 27 items	+.344	-.035	Six-factor, 27-item model
Six-factor model, 27 items, vs. five-factor model, 27 items	Five-factor, 27 items	Six-factor, 27 items	+.011	-.011	Six-factor, 27-item model

Note. Δ values were calculated as target model minus reference model. Positive ΔCFI values and negative ΔRMSEA values indicate comparatively better fit for the target model

(a) the original theoretical model,

(b) a five-factor model derived from the exclusion of the *structure* dimension, and

(c) a six-factor model in which the *ability* domain was differentiated into personal and general components.

These models were subsequently evaluated through CFA.

Confirmatory Factor Analysis

The three models specified in the previous section were estimated and compared using CFA to evaluate their fit to the data. All models were estimated using the WLSMV estimator, appropriate for ordinal data, on an independent sample of 410 participants.

Model fit comparison

The original six-factor theoretical model showed poor fit to the data (see **Table 1**). In contrast, both refined CFA models demonstrated improved fit, indicating that the original structure was not fully supported in the present sample.

Following the comparative strategy described above, changes in CFI and RMSEA were examined to provide a more explicit basis for model selection (see **Table 2**). The comparison showed that both refined 27-item models improved over the original 39-item model. However, when the two refined models were compared directly, the six-factor model showed a small but consistent advantage over the five-factor model.

The empirical five-factor model showed acceptable fit according to conventional criteria. However, the revised six-factor model exhibited consistently better fit across all indices compared to the five-factor model, including improvements in incremental fit indices and reductions in absolute misfit indicators. Although the χ^2/df ratio for the six-factor model (3.03) marginally exceeds the conventional cutoff of 3.00, this index is known to be sensitive to large sample sizes (Byrne, 2016; Kline, 2023). The remaining fit indices were within an acceptable range, supporting the overall adequacy of the model.

Table 3. Standardized loading ranges by dimension in the five- and six-factor CFA models (27-item version)

Factor	Items	Loading range (5F)	Loading range (6F)
Source	1-3, 5	.47-.75	.47-.75
Certainty	6-9	.46-.71	.46-.71
Speed	16, 18, 20	.53-.70	.53-.70
Personal ability	22-26	NS	.65-.83
General ability	27-31	NS	.60-.80
Ability (unified)	22-31	.55-.80	NS
Applicability	34-39	.57-.74	.57-.73

Note. 5F: Five-factor model; 6F: Six-factor model; NS: The factor was not specified in that model

Taken together, the absolute fit indices and the comparative differences in fit indices support the revised six-factor, 27-item model as the best-supported representation of the latent structure among the alternatives tested in this sample.

All standardized factor loadings in the retained 27-item model were positive, statistically significant ($p < .001$), and exceeded the minimum recommended threshold ($\lambda \geq .40$). Loading ranges are presented in **Table 3**. The six-factor model showed a clearer distribution of items within the ability domain when specified as two correlated factors.

Item residual variances (θ) ranged from .306 to .784 across items, indicating moderate levels of measurement error. Lower residual variances were observed for several ability-related items, whereas comparatively higher values appeared in some indicators from the source and certainty dimensions, reflecting lower explained variance at the item level.

Inspection of the standardized residual matrix revealed localized areas of model misfit, with some residuals exceeding commonly recommended thresholds (e.g., $|SR| > 4$). These values suggest the

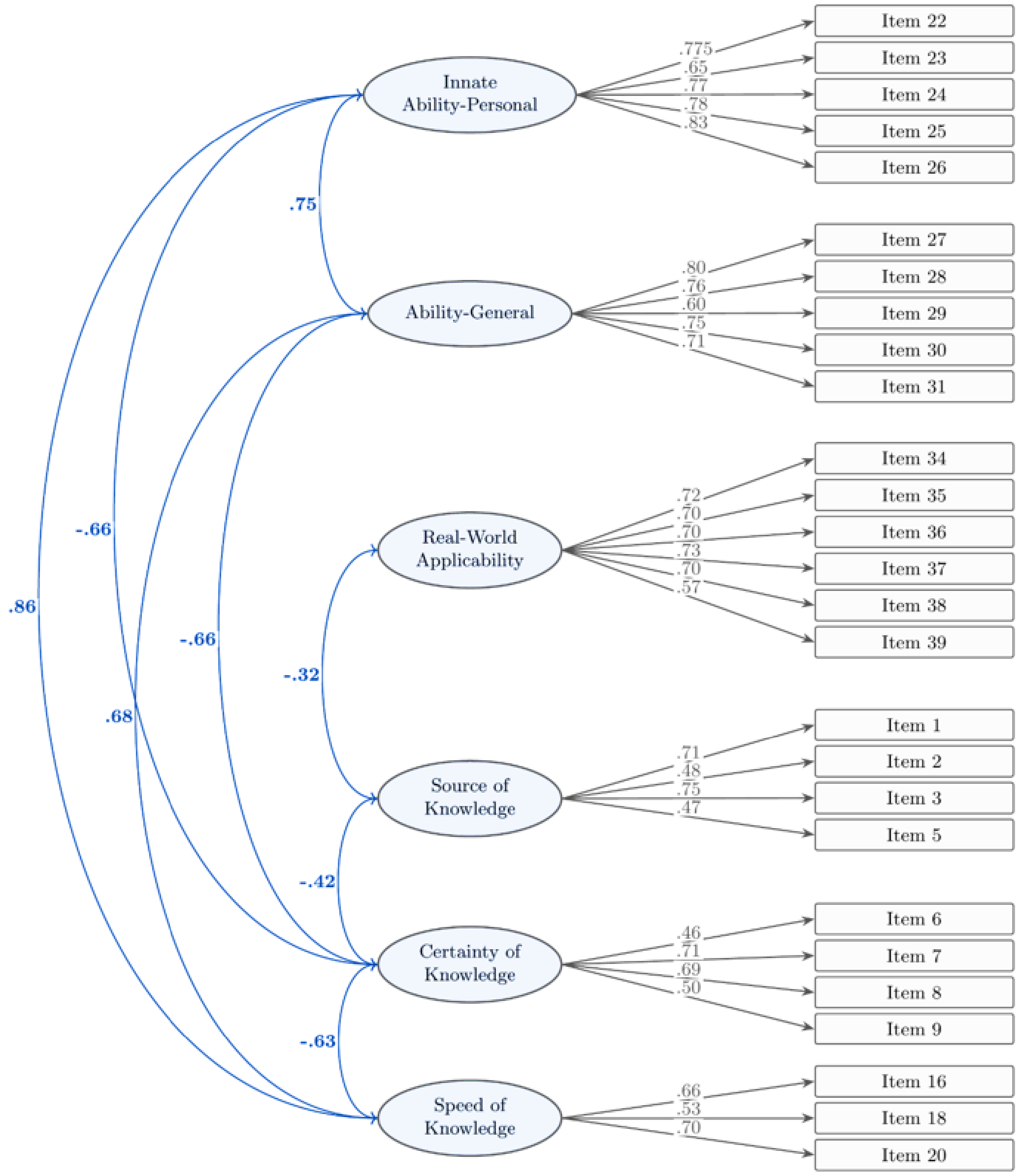

Figure 3. Six-factor confirmatory factor model of the EBSM (standardized factor loadings [λ] are displayed on the arrows from latent variables to observed items, and bidirectional arrows represent latent factor correlations) (the authors' own elaboration)

presence of unexplained covariance between specific item pairs at the local level. No correlated residuals were specified in the final model, in order to preserve its original factorial structure (Kline, 2023).

The confirmatory structure of the retained six-factor model is illustrated in **Figure 3**. The diagram presents standardized factor loadings (λ) and standardized latent correlations (ϕ). For clarity of presentation, residual variances are omitted.

Interfactor correlations (see **Figure 3**) revealed a differentiated pattern of relationships among the latent dimensions. Strong positive associations were observed within the ability domain, particularly between personal and general ability, indicating conceptual coherence within this construct. In contrast, several dimensions exhibited moderate to strong negative correlations, especially between beliefs related to knowledge source, certainty, and speed. This pattern suggests that different epistemic positions may be inversely related in the present sample, although these associations should be interpreted cautiously given the comparatively weaker psychometric performance of some dimensions. Overall, the pattern of interfactor correlations supports a multidimensional structure, although some dimensions exhibited strong associations, particularly between speed and ability-related factors. This pattern indicates that, while the dimensions are distinguishable, certain constructs share substantial variance.

Reliability and validity

Reliability was estimated using standardized ordinal alpha (α), computed from the polychoric correlation

Table 4. Reliability and convergent validity indices for the six-factor model

Dimension	k	α	ω_t	AVE	CR
Source	4	.679	.689	.378	.699
Certainty	4	.668	.679	.358	.683
Speed	3	.667	.670	.405	.669
Personal ability	5	.869	.871	.582	.874
General ability	5	.832	.836	.531	.849
Applicability	6	.840	.841	.474	.843
Total	27	.728	.898	-	-
Criterion	-	≥ .70	≥ .70	≥ .50	≥ .70

Note. *k*: Number of items; α: Standardized ordinal alpha; ω_t: McDonald's omega total; The total row presents overall reliability indices for the 27-item instrument; AVE and CR are not applicable at the total scale level

matrix, and McDonald's omega total (ω_t). Convergent validity was assessed using average variance extracted (AVE) and composite reliability (CR). All estimates for the retained six-factor model are presented in **Table 4**.

Overall reliability for the final 27-item version, estimated on the independent CFA subsample ($n = 410$), was satisfactory ($\alpha = .728$; $\omega_t = .898$), supporting the internal consistency of the adapted instrument.

At the dimensional level, personal ability and general ability demonstrated strong internal consistency and convergent validity, with both AVE and CR exceeding recommended thresholds, indicating cohesive latent structures. Applicability showed acceptable CR, although its AVE was slightly below the recommended cutoff. In contrast, speed approached but did not reach the conventional CR threshold and showed low AVE, indicating that this dimension should be interpreted with caution. Source and certainty yielded AVE values below .50 and CR estimates close to the conventional threshold, indicating relatively lower shared variance among their indicators.

Discriminant validity was evaluated using the Heterotrait-Monotrait (HTMT) ratio of correlations and the Fornell-Larcker criterion (Henseler et al., 2015). HTMT values were generally below the recommended threshold of .85, with the highest value observed between speed and personal ability (HTMT = .848), approaching the upper bound. The Fornell-Larcker criterion indicated limited discriminant validity for this same pair, as the interfactor correlation ($\phi = .859$) exceeded the square root of AVE for both factors (.636 and .763, respectively), suggesting partial overlap between these constructs and reinforcing the need for cautious interpretation of speed as a distinct dimension**.**

Item-level explained variance (R^2), derived from standardized factor loadings, ranged from .216 to .694 across indicators, indicating moderate to adequate item-factor relationships throughout the model.

Additional reliability analysis using a hierarchical model indicated a hierarchical omega of $\omega_h = .622$ and an explained common variance (ECV = .400), supporting the presence of a general factor alongside the multidimensional structure.

Latent factor correlations

Standardized latent correlations (ϕ) in the retained six-factor model indicated moderate to high associations among dimensions. The correlation between personal ability and general ability was high ($\phi = .748$), indicating substantial shared variance between these constructs. Strong associations were also observed between speed and ability-related factors, particularly with personal ability ($\phi = .859$), reflecting close relationships among performance-oriented dimensions.

Across the model, consistent negative latent correlations were observed between certainty and performance-related factors ($\phi = -.63$ to $\phi = -.70$).

For comparison purposes, the five-factor model showed a very high correlation between speed and ability ($\phi = .83$), which supports the subsequent differentiation of the ability dimension in the six-factor solution.

Measurement invariance across educational levels

Measurement invariance of the adapted instrument was examined within the CFA subsample (n = 410), sequentially comparing basic secondary (grades 6-9; n = 258) and upper secondary (grades 10-11; n = 152) students. The configural model provided evidence that the six-factor structure was tenable in both groups (CFI = .906, RMSEA = .069). Metric invariance was supported according to changes in practical fit indices (ΔCFI = -.006; ΔRMSEA = +.001, relative to the configural model), as was scalar invariance (ΔCFI = +.002; ΔRMSEA = -.003, relative to the metric model). These results suggest that factor loadings and item thresholds were sufficiently comparable across educational levels. Strict invariance was not supported (ΔCFI = -.012, relative to the scalar model), indicating that residual equality should not be assumed.

DISCUSSION

The factorial structure obtained in this study departs substantively from the original EBSM. Walker (2007) reported that mathematics-related epistemic beliefs among American university students converged into a single higher-order factor explaining approximately 58% of the common variance, reflecting a coherent, albeit epistemically naïve, structure. The present findings produce a different configuration: hierarchical reliability decomposition of the six-factor model yielded $\omega_h = .622$ and an ECV of .40, indicating that a general epistemic factor is present but accounts for only a moderate portion of shared variance. This result does not support a strongly integrated hierarchical structure; it points instead to a system that is partially unified yet substantively multidimensional. Before interpreting this divergence, a fundamental difference in sampling must be acknowledged. Walker's (2007) validation involved university students whereas the present study focuses on secondary school students. Epistemic beliefs are not static across educational levels, as suggested in prior research: university students have generally had greater exposure to disciplinary discourse, argumentative evaluation, and the negotiation of meaning across domains than secondary students (Hofer & Pintrich, 1997; Schraw et al., 2002). Some degree of epistemic integration may therefore be developmentally expected at the university level, consistent with stage-based models of epistemic development (Hofer & Pintrich, 1997; Perry, 1970), while the moderate hierarchical structure observed here may be consistent with a developmental interpretation. The present study cannot disentangle educational level from instructional context, and distinguishing their relative contributions would require comparative designs beyond the scope of this study.

What the present results do establish is that the moderate general factor (ω_h = .622; ECV = .40) coexists with a highly heterogeneous pattern of interfactor correlations: strong positive associations within the ability domain (ϕ = .748), consistent negative correlations between certainty and performance-related dimensions (ϕ = -.631 to -.664), and near-zero or weak correlations involving applicability (ϕ = -.068 to -.323). The general factor does not resolve these tensions; it coexists with dimensions suggesting a pattern of epistemic beliefs that may be weakly integrated or not fully aligned. A partially unified epistemic system in which the general orientation toward mathematical knowledge does not fully account for the functional oppositions among its dimensions is not equivalent to a naïve but coherent one, and it may carry different implications for instruction and intervention. Whether this configuration is primarily developmental, primarily instructional, or the product of both operating simultaneously is a question this study raises but cannot settle (Hofer & Pintrich, 1997; Muis, 2004). Ernest's (2015) discussion of the unintended outcomes of school mathematics is relevant here, since several beliefs assessed by the adapted instrument—particularly those related to certainty, ability, and applicability—overlap with broader attitudes and assumptions that students may develop through their experience of school mathematics. However, the present study does not test the instructional origins of these beliefs, and this interpretation should therefore be treated as contextual rather than causal.

The Absence of the Structure Dimension

The removal of the *structure of mathematical knowledge* dimension raises questions about an assumption in domain-specific epistemic belief models: that the dimensions proposed by Hofer and Pintrich (1997) and operationalized by Walker (2007) are equally accessible as belief objects across educational contexts. Schommer (1990) conceptualized epistemic beliefs as a loosely interconnected system in which dimensions develop independently, implying that any dimension can in principle emerge in any population given appropriate measurement. The present findings do not fully support that assumption for the case of structure: this dimension did not merely underperform psychometrically; it produced indicators so weak as to suggest the construct may lack a clearly coherent psychological referent in this population (KMO = .49, α = .34, $\bar{r}$ = .11).

One possible theoretical explanation for this lies in what structure, unlike other epistemic dimensions, requires of individuals. Dimensions such as source or certainty can form around direct, everyday epistemic experiences: students encounter authority figures, receive right and wrong answers, and develop beliefs accordingly (Hofer & Pintrich, 1997; Muis, 2004). Structure, by contrast, presupposes a qualitatively different way of engaging with mathematical concepts, one in which students are not only asked to use mathematical ideas operationally, as procedures to be executed, but also to conceive them structurally, as conceptual objects that can be related, transformed, and reorganized (Sfard, 1991). Without that instructional history, students may lack an experiential referent from which structure-related beliefs can cohere (Vygotsky, 1978). This distinction suggests a possible boundary condition in Muis' (2004) domain-specific model: the model assumes that domain-specific epistemic beliefs are more coherent than general ones because domain-specific instruction provides clearer epistemic referents. That assumption may hold for several dimensions, but structure may require a particular kind of domain-specific instruction, one oriented toward conceptual integration, that is not universally present even within mathematics classrooms.

In the present population, that instructional history may not be equally well established. Colombian secondary students may show patterns consistent with developmental accounts of epistemic beliefs, particularly those describing earlier stages characterized by dualistic thinking and reliance on authority (Hofer & Pintrich, 1997; Perry, 1970). Beyond developmental stage, prior research in some Latin American secondary contexts has described mathematics teaching as frequently centered on procedural routines and algorithmic execution, with comparatively limited attention to conceptual integration or relational understanding (Estrella & Chaves, 2023; Montero-Sánchez, 2025). Such conditions, if present, may limit students' opportunities to develop coherent beliefs about the structure of mathematical knowledge, although the present study did not directly measure instructional practices. This interpretation extends the findings reported by Vizcaino Escobar et al. (2015). Whereas that study documented difficulties in the emergence of the structure dimension, the present results suggest that its absence may be associated with instructional and developmental conditions that require direct examination in future research.

Crucially, this outcome was not entirely unpredictable from Walker's (2007) own data. The *structure* dimension was already the most fragile of the seven original factors: it produced the lowest internal consistency estimate (α = .65), the narrowest range of factor loadings (.51-.62), and was the only factor alongside real-world applicability to load negatively (λ = -.65) on the second-order general factor. That negative loading suggests a notable pattern: even among American university students, *structure* operated in the direction opposite to the rest of the epistemic system. Students who held more sophisticated structural beliefs were paradoxically less aligned with the general naïve epistemic orientation that the remaining factors shared. This pattern is consistent with Hofer and Pintrich's (1997) argument that epistemic dimensions do not necessarily develop in parallel or converge toward a unified framework, and with Muis' (2004) observation that certain dimensions require more specific instructional conditions to become psychologically salient. Structure was already functioning as a partially independent epistemic construct in Walker's (2007) sample, one that required specific conditions to remain coherent. Applying the EBSM to younger students in a different educational context may have contributed to the weak performance of the structure factor. However, the present design cannot determine whether this result was due to educational level, instructional context, item interpretation, or a combination of these factors.

This has implications beyond the present validation. If the coherence of structure-related beliefs depends partly on conceptually oriented instruction, then its inclusion in instruments applied across diverse educational contexts cannot be assumed without first establishing that those conditions are present (Dorsah et al., 2025; Říčan & Pešout, 2022; Schraw et al., 2002). The finding reported here is not that the EBSM structure items are poor measures; it is that the construct they measure may be educationally contingent in ways that existing multidimensional models of epistemic beliefs have not fully theorized (Hofer & Pintrich, 1997; Muis, 2004). Future work should examine whether targeted exposure to conceptually integrated mathematics instruction produces corresponding changes in the coherence of structure-related beliefs, a question that would directly test the instructional contingency hypothesis proposed here. Cognitive interviewing of item interpretation across educational levels would additionally clarify whether the absence of this dimension reflects instructional conditions, developmental stage, or item-level interpretive variation.

Taken together, these findings indicate that the absence of the structure dimension should not be interpreted solely as a psychometric deficiency of the instrument. Rather, it may also reflect context-dependent conditions that constrain the emergence of this construct as a coherent belief dimension. Specifically, the absence of a coherent structure factor may be related to students' developmental stage, their opportunities to engage with mathematics as a structured and interconnected system, or item-level interpretation. However, because developmental progression and instructional practices were not directly measured, these explanations should be treated as plausible hypotheses rather than direct empirical conclusions.

Under these conditions, students may have limited opportunities to engage with mathematics as a structured and interconnected system, which may constrain the formation of structure-related epistemic beliefs. In this sense, the present findings raise the possibility that certain epistemic dimensions are not universally accessible across populations but instead may depend on the interaction between cognitive development and the epistemic affordances of the learning environment.

The Peripheral Status of the Speed Dimension

The reduction of *speed of knowledge* to a three-item core requires careful interpretation. Psychometrically, the dimension shows the weakest profile of the retained factors: the lowest AVE (.405), the fewest items, and the highest proportion of cross-loadings in the EFA. In isolation, these indicators might suggest marginal factorial status. The interfactor correlations suggest a different pattern. Speed is strongly associated with personal ability (ϕ = .859) and moderately associated with general ability (ϕ = .676), while showing consistent negative correlations with certainty (ϕ = -.631). This does not correspond to the pattern of a fully independent dimension; it is the pattern of a dimension that is relationally embedded but psychometrically underdeveloped and should therefore be interpreted with caution.

This combination has a specific interpretation in the present context. Students in this sample may not clearly distinguish between believing that mathematics requires speed and believing that mathematical ability is a personal or general capacity, partly because the retained items link rapid problem solving, understanding, need for help, time, effort, and perceived competence. The high correlation between speed and personal ability suggests that being fast and being capable appear to be closely associated in students' perceptions, possibly reflecting performance-oriented evaluation contexts

(Hofer & Pintrich, 2002; Muis, 2004). When speed functions as a primary signal of competence in certain instructional contexts, students develop speed beliefs not as an independent epistemic stance about the nature of mathematical knowledge, but as an extension of their beliefs about mathematical ability.

The reduction of the dimension to three items does not necessarily reflect the absence of speed-related beliefs in this population but rather may reflect a lack of differentiation from ability beliefs. Walker (2007) herself noted that general attitudes toward mathematics may mask more subtle distinctions among dimensions. It is also compatible with broader theoretical distinctions between beliefs about knowledge and beliefs about learning, since the retained speed items refer primarily to rapid understanding, problem solving, and need for help, rather than to the structure of mathematical knowledge itself. In the present sample, this masking may operate specifically at the boundary between speed and ability, producing a reduced but relationally active dimension whose weak psychometric profile should be treated as a limitation requiring further validation rather than as a fully established substantive finding. Psychometric evidence across cultural contexts has also reported that speed-related items consistently load weakly and unstably (Dorsah et al., 2025). Speed should therefore not be interpreted as a fully consolidated independent dimension; its reduced psychometric profile may reflect the absorption of speed-related content into the broader belief structure surrounding mathematical competence.

The Differentiation of Ability Beliefs

The differentiation of ability-related beliefs into personal ability and general ability constitutes one of the clearest structural findings in this sample. Walker (2007) originally proposed this distinction and found that *innate ability-general* exhibited stronger external validity, showing more robust associations with fixed theories of intelligence than its personal counterpart. The present findings provide support for retaining this distinction in the adapted version, in a context different from the one in which the instrument was originally developed.

This differentiation is theoretically consistent with broader models of belief systems related to competence. Bandura (1997) distinguishes between efficacy judgments, that is, beliefs about one's own capacity to perform specific tasks, and more general beliefs about the underlying abilities that support performance. This distinction remains consistent with contemporary social-cognitive accounts of motivation, which continue to treat self-efficacy as a task- and context-specific judgment of capability (Schunk & DiBenedetto, 2020). In this sense, personal ability beliefs may reflect self-evaluations of competence, whereas general ability beliefs may reflect broader conceptions about the nature and distribution of mathematical talent.

The moderate to strong correlation between these factors suggests that they are related but not redundant constructs, supporting their specification as distinct but related dimensions within the adapted model. Their coexistence suggests that students may simultaneously hold beliefs about their own mathematical competence and broader assumptions about whether such competence is inherently determined, a configuration that may have implications for how they engage with mathematical learning and persist in challenging tasks.

A Partially Integrated Epistemic System

The results indicate that epistemic beliefs about mathematics in this population operate as a partially integrated system rather than a fully coherent structure. The negative associations between *certainty of knowledge* and performance-related dimensions replicate a well-established finding: absolutist conceptions of knowledge co-occur with passive learning strategies and limited conceptual engagement (Hofer & Pintrich, 1997; Muis, 2004; Schiefer et al., 2022). In the context of the present findings, however, this pattern carries a more specific implication. Students who conceive mathematical knowledge as fixed and externally validated may be simultaneously less likely to perceive themselves as active agents in its construction, as these beliefs may co-occur and may be consistent with the same instructional logic: knowledge is transmitted, not built, and the student's role is receptive, not generative (Hofer & Pintrich, 1997; Muis, 2004).

In this configuration, these dimensions do not form the differentiated-but-integrated structure that Hofer and Pintrich (1997) theorized or the coherent naïve cluster that Walker (2007) observed. The presence of a general epistemic factor (ω_h = .622) indicates that beliefs are not entirely unrelated, but the heterogeneous pattern of interfactor correlations suggests a system in which epistemic positions may be functionally opposed rather than merely differentiated. Addressing this configuration may require more than exposing students to sophisticated epistemic models; it may also involve creating the instructional conditions that make certain epistemic categories, namely structure, conceptual connection, and active knowledge construction, experientially available as objects of belief. The coexistence of high-certainty beliefs, beliefs about innately distributed mathematical ability, and the absence of a coherent structure factor should not be reduced to an early developmental explanation alone; it may also be consistent with a configuration of beliefs that could be less compatible with inquiry-based or reform-oriented instruction.

These findings extend previous Latin American validation studies, such as Vizcaino Escobar et al. (2015), which interpreted factorial reorganization primarily as evidence of contextual sensitivity. The present results identify a specific structural consequence of that sensitivity: not a reduced or rearranged factor model, but one in which a moderate general factor coexists with dimensions that may remain functionally opposed under conditions that do not promote conceptual integration. This distinction may inform how mathematics educators and researchers interpret epistemic belief data in Latin American contexts and how they design or evaluate instructional interventions.

Limitations and Future Research

Several considerations should guide the interpretation of these findings. The study was based on a large sample of 817 Colombian secondary students from public schools, covering grades 6-11 (ages 10-20), which provides a solid empirical basis for the psychometric evaluation of the adapted EBSM. However, because the sample was obtained through non-probabilistic purposive sampling within specific public-school contexts, the generalizability of the findings beyond the institutional contexts included in the study remains limited. Future studies should therefore examine the stability of the proposed structure across additional Colombian regions, school sectors, and educational settings. In addition, the lack of external criterion variables constrains the evaluation of the instrument's nomological network, and the cross-sectional design prevents conclusions about developmental progression. The absence of direct measures of classroom instruction, curriculum exposure, and item interpretation also means that possible explanations for the non-retention of the structure dimension should be treated as theoretically plausible hypotheses rather than empirical conclusions. Finally, although speed was retained in the six-factor model, its weaker psychometric profile and high association with personal ability indicate that this dimension should be interpreted with appropriate caution. Future research should address these issues by examining the coherence of structure-related beliefs, refining or expanding speed-related items, incorporating cognitive interviews to examine item interpretation, and evaluating criterion-related validity using relevant educational outcomes.

CONCLUSIONS

The adapted 27-item instrument showed acceptable psychometric properties and a confirmatory factor model that fit the data substantially better than the original 39-item model. Beyond its measurement performance, the findings provide insight into the organization of epistemic beliefs in this sample. Rather than indicating a fully integrated belief system, the results suggest that epistemic beliefs about mathematics were only partially integrated. This interpretation is supported by the non-retention of the structure dimension, the close association between speed and ability-related constructs, and the negative relationships between certainty-oriented and performance-related dimensions. Although a general epistemic factor was present, it did not fully account for the pattern of relationships among dimensions.

These results suggest that the EBSM may not function identically across educational levels and local educational settings. Because the instrument was originally developed with university students, its application to secondary populations may require not only psychometric adaptation but also consideration of whether the epistemic dimensions it assumes are equally salient and measurable across educational settings. In this sense, the non-retention of the structure dimension should be interpreted as a context-sensitive psychometric finding rather than as evidence of a direct cultural, developmental, or instructional cause. Future research should examine whether structure-related beliefs vary according to curriculum exposure, instructional practices, and developmental level.

From an applied perspective, the adapted instrument may provide a useful starting point for researchers and educators interested in examining epistemic beliefs in Colombian secondary education. However, further validation is needed before broad application. At this stage, the instrument is best understood as suitable for research contexts in which its limitations are explicitly acknowledged, particularly when exploring relationships between epistemic beliefs, learning strategies, and academic performance. Overall, the results provide evidence for the 27-item EBSM as a psychometrically supported measure for this Colombian secondary-school sample, while underscoring the importance of further studies examining its stability, measurement invariance, and external validity across broader and more diverse educational contexts.

Author contributions: ESM-G: conceptualization, formal analysis, methodology, writing – original draft, writing – review & editing; **EAP-R:** data curation, validation, writing – review & editing; **RD-R:** methodology, supervision, writing – review & editing. All authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: This study was approved by the Ethics Committee of Universidad de Investigación y Desarrollo, Bucaramanga, Colombia. Informed consent was obtained from parents or legal guardians of all participants. Student assent was also obtained prior to participation.

AI statement: OpenAI ChatGPT (GPT-5.6 Sol) was used during manuscript revision to assist with wording, clarity, and stylistic refinement in selected passages. All AI-assisted output was reviewed and verified by the authors, who take full responsibility for the final manuscript.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Alkış Küçükaydin, M., & Ayaz, E. (2024). Validation of the scientific reasoning competencies instrument: Relationships with epistemological beliefs and analytical thinking. *International Journal of Science and Mathematics Education, 23*, 343-363. https://doi.org/10.1007/s10763-024-10482-2

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct (2002, amended effective June 1, 2010, and January 1, 2017). *APA*. https://www.apa.org/ethics/code

Arias Ortiz, E., Giambruno, C., Morduchowicz, A., & Pineda, B. (2024). *The state of education in Latin America and the Caribbean 2023*. Inter-American Development Bank. https://doi.org/10.18235/0005515

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Boaler, J. (2022). *Mathematical mindsets: Unleashing students' potential through creative mathematics, inspiring messages and innovative teaching* (2nd ed.). Jossey-Bass.

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

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (3rd ed.). Routledge. https://doi.org/10.4324/9781315757421

Cresswell, C., & Speelman, C. P. (2020). Does mathematics training lead to better logical thinking and reasoning? A cross-sectional assessment from students to professors. *PLoS ONE, 15*(7), Article e0236153. https://doi.org/10.1371/journal.pone.0236153

Dorsah, P., Amedeker, M. K., & Ngman-Wara, E. I. (2025). Exploring the factor structure of the epistemic beliefs inventory. *International Journal of Didactical Studies, 6*(2), Article 33004. https://doi.org/10.33902/ijods.202533004

Ernest, P. (1989). The knowledge, beliefs and attitudes of the mathematics teacher: A model. *Journal of Education for Teaching, 15*(1), 13-33. https://doi.org/10.1080/0260747890150102

Ernest, P. (2015). The social outcomes of learning mathematics: Standard, unintended, or visionary? *International Journal of Education in Mathematics, Science and Technology, 3*(3), 187-192. https://doi.org/10.18404/ijemst.29471

Estrella, S., & Chaves, E. (2023). Balance y perspectivas de la estadística y la probabilidad en la educación matemática [Balance and perspectives of statistics and probability in mathematics education]. In P. Scott, Y. Morales, & Á. Ruiz (Eds.), *Educación matemática en las Américas 2023: Trabajos invitados de la XVI CIAEM* (vol. 1, pp. 302-311). Comité Interamericano de Educación Matemática.

Hambleton, R. K., & Zenisky, A. L. (2011). Translating and adapting tests for cross-cultural assessments. In D. Matsumoto, & F. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 46-70). Cambridge University Press. https://doi.org/10.1017/CBO9780511779381.004

Henseler, J., Ringle, C., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*, 115-135. https://doi.org/10.1007/s11747-014-0403-8

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140. https://doi.org/10.3102/00346543067001088

Hofer, B. K., & Pintrich, P. R. (Eds.). (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410604316

Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human-Computer Interaction, 32*(1), 51-62. https://doi.org/10.1080/10447318.2015.1087664

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

International Test Commission. (2018). The ITC guidelines for translating and adapting tests (second edition). *International Journal of Testing, 18*(2), 101-134. https://doi.org/10.1080/15305058.2017.1398166

Jablonski, S. (2025). *Modelling, argumentation and problem-solving in the context of outdoor mathematics: Results of the MAP study*. Springer. https://doi.org/10.1007/978-3-658-47751-6

Kelly, M. (2020). Epistemology, epistemic belief, personal epistemology, and epistemics: A review of concepts as they impact information behavior research. *Journal of the Association for Information Science and Technology, 71*(11), 1294-1306. https://doi.org/10.1002/asi.24422

Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.

Montero-Sánchez, N. (2025). Algorithmic thinking in secondary mathematics education: Challenges, strategies, and opportunities. *Sciencevolution, 4*(3), 47-60. https://doi.org/10.61325/ser.v4i3.206

Muis, K. R. (2004). Personal epistemology and mathematics: A critical review and synthesis of research. *Review of Educational Research, 74*(3), 317-377. https://doi.org/10.3102/00346543074003317

OECD. (2023). *PISA 2022 results (volume II): Learning during – and from – disruption*. OECD Publishing. https://doi.org/10.1787/a97db61c-en

Ondaro, J. E. V., & Dagoc, J. C., Jr. (2025). Students' self-efficacy, epistemological beliefs and mathematics performance. *International Journal of Research and Innovation in Social Science, 9*(3S), 4577-4589. https://doi.org/10.47772/IJRISS.2025.903SEDU0329

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. Holt, Rinehart and Winston.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*, 539-569. https://doi.org/10.1146/annurev-psych-120710-100452

Říčan, J., & Pešout, O. (2022). Analýza sebeposuzovacích dotazníků zjišťujících epistemická přesvědčení žáků druhého stupně ZŠ ve vyučovacím předmětu dějepis [Analysis of self-assessment questionnaires assessing epistemic beliefs of second-grade elementary school students in the subject of history]. *Pedagogika, 72*(4), 455-490. https://doi.org/10.14712/23362189.2022.2266

Rincón-Álvarez, G. A., Hernández-Suárez, C. A., Prada-Núñez, R., Solano-Pinto, N., & Fernández-Cézar, R. (2022). Cuestionario de creencias sobre las matemáticas: Propiedades psicométricas [Questionnaire on beliefs about mathematics: Psychometric properties]. *Educación y Ciudad*, (43), 215-236. https://doi.org/10.36737/01230425.n43.2022.2687

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Sandoval, W. A., Greene, J. A., & Bråten, I. (2016). Understanding and promoting thinking about knowledge: Origins, issues, and future directions of research on epistemic cognition. *Review of Research in Education, 40*, 457-496. https://doi.org/10.3102/0091732X16669319

Schiefer, J., Edelsbrunner, P. A., Bernholt, A., Kampa, N., & Nehring, A. (2022). Epistemic beliefs in science—A systematic integration of evidence from multiple studies. *Educational Psychology Review, 34*, 1541-1575. https://doi.org/10.1007/s10648-022-09661-w

Schoenfeld, A. H. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education, 20*(4), 338-355. https://doi.org/10.2307/749440

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498-504. https://doi.org/10.1037/0022-0663.82.3.498

Schraw, G., Bendixen, L. D., & Dunkle, M. E. (2002). Development and validation of the epistemic belief inventory (EBI). In B. K. Hofer, & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 261-276). Routledge.

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology, 60*, Article 101832. https://doi.org/10.1016/j.cedpsych.2019.101832

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1-36. https://doi.org/10.1007/BF00302715

Vizcaino Escobar, A. E., Manzano Mier, M., & Casas Cardoso, G. (2015). Validez de constructo y confiabilidad del cuestionario de creencias epistemológicas sobre la matemática en alumnos de secundaria básica [Construct validity and reliability of the questionnaire on epistemological beliefs about mathematics in junior high school students]. *Revista Colombiana de Psicología, 24*(2), 301-316. https://doi.org/10.15446/rcp.v24n2.43974

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Walker, D. L. W. (2007). *The development and construct validation of the epistemological beliefs survey for mathematics* [Unpublished PhD thesis]. Oklahoma State University.

Watkins, M. W. (2020). *A step-by-step guide to exploratory factor analysis with R and RStudio*. Routledge. https://doi.org/10.4324/9781003120001

World Bank. (2024). *World development report 2024: The middle-income trap*. World Bank. https://doi.org/10.1596/978-1-4648-2078-6

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191-2194. https://doi.org/10.1001/jama.2013.281053

APPENDIX A: FINAL 27-ITEM VERSION OF THE EBSM (SPANISH VERSION)

Instrucciones. Indique el grado en que está de acuerdo con cada afirmación utilizando una escala tipo Likert de 1 (Totalmente en desacuerdo) a 5 (Totalmente de acuerdo).

Nota. Los números en superíndice indican la numeración original de los ítems en la versión de 39 ítems.

Table A1. Final 27-item version of the EBSM (Spanish version)

No	Ítem
Fuente del conocimiento	
1	Aprender matemáticas depende de lo bueno que sea el profesor(1).
2	Aprendo mejor matemáticas cuando el profesor desarrolla ejemplos en sus explicaciones(2).
3	La calidad de las clases de matemáticas depende completamente del profesor(3).
4	Para resolver problemas de matemáticas, es necesario aprender el procedimiento correcto(5).
Certeza del conocimiento	
5	En matemáticas, una respuesta solo puede ser correcta o incorrecta(6).
6	En las clases de matemáticas no hay espacio para la creatividad(7).
7	Todos los profesores de matemáticas siempre dan la misma respuesta cuando se les pregunta sobre un tema(8).
8	En matemáticas, lo que es verdadero no cambia con el tiempo(9).
Velocidad de adquisición del conocimiento	
9	Cuando se introduce un tema nuevo en matemáticas, la mayoría lo entiende rápidamente(16).
10	Si no puedo resolver un problema de matemáticas en poco tiempo, necesitaré ayuda para resolverlo(18).
11	Si entiendo bien un tema de matemáticas, debería poder resolver los problemas rápidamente(20).
Habilidad personal	
12	Si me lo propongo, podría aprender cosas más avanzadas en matemáticas(22).
13	Para mí, las matemáticas son difíciles de aprender: por más que me esfuerce, siento que nunca llegaré a entenderlas(23).
14	Si las matemáticas fueran fáciles, no tendría que dedicar tanto tiempo a realizar las tareas(24).
15	Me desanima tener que estudiar mucho para entender un problema de matemáticas(25).
16	Creo que no nací con talento para las matemáticas, aunque tengo habilidades en otras áreas(26).
Habilidad general	
17	Las matemáticas no son para todo el mundo; algunas personas nacen con esta habilidad y otras no(27).
18	Ser bueno en matemáticas es algo con lo que se nace(28).
19	Los estudiantes con habilidades para las matemáticas no necesitan resolver muchos ejercicios porque comprenden los contenidos rápidamente(29).
20	Para ser bueno en matemáticas se requiere haber nacido con esa habilidad(30).
21	Las personas saben desde niños si son buenas o no para las matemáticas(31).
Aplicabilidad en el mundo real	
22	Saber matemáticas es importante para algunas profesiones, pero no para todas(34).
23	Necesito aprender matemáticas para ser mejor profesional en el futuro(35).
24	Puedo aplicar en otras materias lo que aprendo en matemáticas(36).
25	Es fácil saber cómo aplicar las matemáticas en situaciones de la vida diaria(37).
26	Las matemáticas son importantes porque se aplican en muchas profesiones(38).
27	Las matemáticas nos ayudan a entender mejor el mundo en el que vivimos(39).

APPENDIX B: SUPPLEMENTARY TABLES

Table B1. Standardized factor loadings for the six-factor CFA model (27-item version, n = 410)

	Factor/item	λ
Source of knowledge	CEM03	.746
	CEM01	.712
	CEM02	.477
	CEM05	.468
Certainty of knowledge	CEM07	.713
	CEM08	.682
	CEM09	.493
	CEM06	.465
Speed of knowledge	CEM20	.700
	CEM16	.664
	CEM18	.533
Personal ability	CEM26	.833
	CEM25	.776
	CEM22	.775
	CEM24	.767
	CEM23	.651
General ability	CEM27	.799
	CEM28	.758
	CEM30	.753
	CEM31	.712
	CEM29	.604
Real-world applicability	CEM37	.734
	CEM34	.727
	CEM36	.700
	CEM35	.696
	CEM38	.695
	CEM39	.568

Note. Items are ordered by loading magnitude within each factor. All loadings are statistically significant ($p < .001$)

Table B2. Comparison of standardized loadings for ability-related items in the five- and six-factor models

Item	5F	6F	
	Ability	Personal ability	General ability
CEM26	.797	.833	
CEM25	.748	.776	
CEM22	.746	.775	
CEM24	.734	.767	
CEM23	.623	.651	
CEM27	.728		.799
CEM28	.700		.758
CEM30	.701		.753
CEM31	.657		.712
CEM29	.553		.604
Range	.553-.797	.651-.833	.604-.799

Note. In the five-factor model, all ten items load on a single unified ability factor. In the six-factor model, items CEM22–CEM26 load on personal ability and items CEM27-CEM31 load on general ability. Blank cells indicate that the item does not belong to that factor in the six-factor solution. All loadings are statistically significant ($p < .001$)

Table B3. Standardized interfactor correlations (φ) for the five- and six-factor CFA models: Five-factor model

	Source of knowledge	Certainty of knowledge	Speed of knowledge	Ability (unified)	Real-world applicability
Source	1.000				
Certainty	-.417	1.000			
Speed	.351	-.631	1.000		
Ability	.392	-.704	.830	1.000	
Applicability	-.323	.248	-.279	-.129	1.000

Table B4. Standardized interfactor correlations (φ) for the five- and six-factor CFA models: Six-factor model

	Source of knowledge	Certainty of knowledge	Speed of knowledge	Personal ability	General ability	Real-world applicability
Source	1.000					
Certainty	-.417	1.000				
Speed	.351	-.631	1.000			
Personal ability	.361	-.664	.859	1.000		
General ability	.378	-.661	.676	.748	1.000	
Applicability	-.323	.248	-.280	-.068	-.182	1.000

Note. The decomposition of the unified ability factor reveals differentiated associations with speed: personal ability ($\varphi = .859$) is more strongly associated with speed than general ability ($\varphi = .676$), supporting the theoretical distinction between the two subdimensions