



# Teachers' Perceptions of Argumentation in Citizenship Education: Psychometric Validation of the AASES Instrument and Mediation Analysis of Sociodemographic Variables Using SEM

Delfín Ortega-Sánchez<sup>1</sup> · Carlos Pérez-González<sup>2</sup>

Received: 18 April 2025 / Accepted: 15 August 2025

© The Author(s) 2025

## Abstract

This study aims to analyze the perceptions of Spanish secondary school teachers in the fields of social and experimental sciences ( $n=215$ ) regarding the formative relevance of argumentative competence in the context of citizenship education. It also seeks to provide a psychometric instrument supported by robust empirical evidence of reliability and validity to achieve this goal. Based on a non-experimental, cross-sectional design, the AASES (Assessment of Argumentation in Social and Experimental Sciences) instrument,—developed ad hoc,—is administered to identify the potential relationship of sociodemographic factors with the theoretical sub-constructs it comprises, the statistical association between age, origin, and gender, and the mediating role of gender in the relationship between age/place of origin and the defined latent factors, and the existence of significant differences among sociodemographic groups. The results indicate that, although a statistical association between age and gender was observed, the mediating role of gender in the relationship between age, origin, and the latent factors cannot be confirmed. SEM analyses showed that none of the sociodemographic predictors (age, origin, and gender) exert statistically significant direct or mediating effects on the latent factors. Furthermore, the comparative analyses complement the model by indicating that perceptions vary moderately by gender and age, even though these variables do not explain the latent factors in the SEM. Indeed, the analysis of variance revealed significant differences in the Critical and Ethical Skills dimension based on these two factors, with higher scores among male teachers and in the oldest age group, as well as increasing trends, with age, in the perceived importance of argumentation for the development and acquisition of critical thinking skills, informed decision-making, and ethical discussion. The findings highlight the need to incorporate specific spaces within the curriculum and teacher education programs to foster argumentative competencies and informed engagement with controversial

---

Extended author information available on the last page of the article

socio-scientific issues, taking into account sociodemographic variables as influential factors in the educational process.

**Keywords** Argumentation · Citizenship education · Teacher perceptions · Psychometric validation · SEM (Structural equation modeling)

## 1 Introduction

Argumentative competence is commonly defined as a complex set of skills encompassing metacognitive (reflection on one's own learning), metastrategic (mastery of specific strategies and the ability to adapt them to context), and epistemological abilities (understanding of one's own knowledge, its foundations, and limitations) (Kuhn et al. 2013; Rapanta et al. 2013). This set of skills is thus aimed at the practice of constructing sound reasoning and fostering reflective dialogue based on the evaluation of evidence, the adaptation of arguments, and the constructive management of dissent, rather than the memorization of logical errors or the identification of a potential catalogue of fallacies in this domain (Blair 2023). Nevertheless, van Eemeren and Garsen (2023) justify the relevance of fallacies as central elements for understanding and improving argumentation, addressing them as violations of the rules of reasonable discussion that hinder the fair resolution of differences of opinion.

This competency therefore entails the ability to select, construct, and organize arguments within a specific context, to demonstrate critical reasoning skills, and to effectively communicate well-founded ideas (Rapanta et al. 2013). Accordingly, the teaching of argumentation entails instruction in the linguistic and structural features of argumentative discourse, aiming to facilitate mastery of signs and symbols—expressed through language—and the cognitive and communicative coherence of ideas. Argumentative competence thus functions as an activator of the skills required for the exchange of arguments and the effective exercise of commitment, responsibility, and rationality in social decision-making within problematic or scientifically controversial contexts.

The scientific literature positions argumentative competence as a structural axis of citizenship education due to its contributions to the promotion of critical thinking, civic engagement, and the willingness to engage in dialogue and deliberation on social, scientific, and political issues. Specific educational programs—such as those that integrate argumentation in socio-scientific contexts and foster the development of social competencies—have shown that active participation enhances the skillset associated with argumentative competence (Evagorou et al. 2023; Fuentes-Moreno et al. 2020; Sengul 2019). Consequently, this competence constitutes one of the foundational elements of education for democratic citizenship, both in its monologic dimension (internal discourse) and in its dialogic dimension, when developed through debate.

In this regard, with the aim of examining the convergences and dialectical influences between education and practices in critical thinking and the most recent contributions from argumentation theory, Olmos (2022) concludes that, despite sharing pedagogical goals, these fields have not always fully converged. The study also high-

lights the essential potential of interactive argumentative practices for the rational and comprehensive development of critical thinking, in line with its contemporary social and educational purposes.

### 1.1 Teaching and Learning to Argue from an Interdisciplinary Perspective

The production of argumentative texts and the teaching of argumentation have garnered increasing interest in educational research, with findings converging on the recognition of argumentation as a key competence for the development of critical thinking and the construction of knowledge. However, the complexity of teaching argumentation lies in the lack of consensus regarding what argumentative competence precisely entails and how it should be assessed. In this regard, Pérez-Echeverría et al. (2016) propose a multifocal analysis of argumentation: structural-formal (logical), dialogic, and socio-constructivist (integrating metacognitive and epistemological dimensions). Shi (2020) advocates for a more precise assessment of argumentative competences, suggesting that written “constructed dialogues” may more accurately reflect students’ actual argumentative quality compared to traditional essays. Likewise, the search for analytical and assessment tools capable of capturing the complexity of argumentative discourse has been addressed by Martins and Macagno (2022), who designed a coding tool for dialogues aimed at identifying both individual and collective purposes, along with argumentative moves that foster the construction of scientific knowledge.

In this vein, Asterhan and Schwarz (2016), through the Argumentation for Learning (AFL) model, emphasize the need to move beyond the act of arguing to consider the prior conditions, discourse features, and learning outcomes as interconnected explanatory dimensions. Similarly, Nussbaum (2020) developed the framework of Critical Integrative Argumentation (CIA), designed to promote the formulation of integrative refutations that take into account the multiplicity of factors involved in a problem. Along these lines, Spagnolo (2024) demonstrated that argumentative competence cannot be fully developed without addressing both teachers’ and students’ beliefs and conceptions about what it means “to argue.” This finding underscores the importance of aligning pedagogical approaches with the pre-existing representations of the various educational actors in order to generate stable and sustainable changes in classroom practices.

Argumentative dialogue has been extensively studied as a fundamental mechanism for developing critical and collaborative thinking (Rapanta and Macagno 2016), as evidenced by Macagno’s (2016) study on the stimulation of peer oral discussions, which fostered a more sophisticated and critical use of evidence compared to individual writing. Matos (2021) further expanded on this idea by arguing that collaborative writing can serve as an effective bridge between oral discussion and individual production of argumentative texts, as it helps consolidate and transfer debated content into other tasks.

Iordanou and Constantinou (2015) reported on the reflective effectiveness of dialogic debate in fostering skills for the rigorous use of evidence. Their research demonstrated that students who engaged in discussions on controversial topics, supported by metacognitive activities, achieved better outcomes than those who merely studied the content. Along these lines, Vicente et al. (2024) explored the practical effects

on argumentation and decision-making of a role-playing activity focused on animal experimentation, introduced as a socio-scientific issue to prospective Early Childhood Education teachers. This methodological approach, grounded in the discussion of ethical and scientific controversies, proved effective in reshaping students' initial positions and balancing emotional and rational arguments. These findings align with the recognized importance of deliberative practice and feedback aimed at the construction and refutation of arguments. Argumentative competence—particularly in its behavioral dimension (the effective practice of arguing)—is closely related to a deeper understanding of specific content. Indeed, although there appears to be no direct relationship between theoretical knowledge about argumentation or attitudes toward it and academic performance, research findings indicate that those who actively engage in argumentative practice achieve greater knowledge acquisition in the area of study (Valero et al. 2022).

There is, in fact, evidence of improved learning outcomes when explicit instruction is implemented to integrate conflicting information within interventions that combine deliberative dialogues and argumentative synthesis writing on socio-scientific issues (Casado-Ledesma et al. 2021). In this regard, Rapanta and Felton (2022), through a review of 143 studies, confirm that deliberative tasks involving genuine disagreement can generate more sophisticated and developed arguments. Similarly, Morawski and Budke (2019) demonstrated the benefits of specific and structured feedback in enhancing the complexity and argumentative quality of students' texts in Geography classrooms, particularly through peer feedback (both oral and digital). Structured reflection thus plays a crucial role in argumentative self-regulation. On this point, Iordanou (2022) argues that the combination of dialogic practice and metacognitive reflection enhances students' ability to use evidence strategically and, even, to challenge their own assumptions.

From this perspective, specific interdisciplinary curricula have been proposed, grounded in dialogic argumentation (intense dialogic interaction) and subsequent reflective activities. Rapanta (2021) and Iordanou and Rapanta (2021) demonstrate that a method such as *Argue With Me* (AWM) can, indeed, be implemented across various disciplines (Language Arts, Science, History, Citizenship Education), confirming both metacognitive and epistemic benefits, as well as positive impacts on the transfer of argumentative skills to different topics and contexts. In this regard, the review by Bubikova-Moan and Sandvik (2022) on argumentation in early childhood suggests that the initial years of schooling constitute a fertile ground for this area. This conclusion aligns with the “learning-to-learn” model proposed by Kuhn and Halpern (2022), in which argumentation is conceived as a skill that can be cultivated from an early age, particularly within the domain of civic education, through the development and application of emerging argumentative capacities.

With respect to teacher education, the study by Altun and Ozsevgec (2025) demonstrated that implementing a training module with explicit oral and written argumentation activities enables prospective science teachers to develop greater didactic self-confidence and transferable argumentative skills for the classroom. Similarly, Ucar-Longford et al. (2024) showed the effectiveness of comprehensive online support in enhancing argumentation skills among pre-service teachers. Both studies underscore the significance of metacognition as a key factor in promoting self-reg-

ulation and critical reflection, aimed at fostering the habitual practice of argumentation. In this domain, the studies by Mallahi (2024) and Wang (2025), conducted in EFL (English as a Foreign Language) contexts in Iran and China respectively, highlight the need for pedagogical scaffolds and technological supports to overcome linguistic and motivational barriers that hinder argumentative writing: while Mallahi observed a predominant use of general strategies that compromise textual coherence, Wang found that online collaborative environments can enhance the formulation of counterarguments and the strategic use of evidence.

Argumentative competence functions as an enabler of the skills required for the exchange of arguments and the effective exercise of commitment, responsibility, and rationality in social decision-making within democratic contexts. In this regard, Canals' (2007) research—aimed at fostering argumentative processes in secondary education classrooms and assessing their contributions to the development of more rational, structured, nuanced, and socially engaged knowledge—already provided evidence of the valuable role that teaching argumentative skills plays in cultivating social and critical thinking. It also underscored the need for its curricular integration to promote active citizenship committed to addressing social issues. Nevertheless, the study revealed persistent limitations in the progression of critical thinking relative to students' initial stances. Canals' instructional proposal remains useful in current teaching and learning processes in the field of argumentation in social sciences, particularly in supporting the construction and visualization of arguments in social learning, organizing ideas effectively, and connecting premises, counterarguments, and conclusions in a logically coherent manner within educational settings.

In this context, the present study aims to identify, through both descriptive and inferential analysis, the perceptions of Spanish social science and science teachers ( $n=215$ ) regarding the relevance of teaching argumentation, and to provide an instrument with sufficient empirical evidence of reliability and validity to support this goal. In line with the theoretical model, two clearly differentiated latent factors are posited. Factor 1 – *Critical and Ethical Skills* (items Q01–Q05) captures perceptions regarding the usefulness of argumentation in promoting critical reasoning, informed decision-making, and ethical deliberation in Social and Experimental Sciences. Factor 2 – *Democratic Participation and Conflict Resolution* (items Q06–Q10) reflects the value attributed to argumentation as a tool to foster active citizenship, democratic participation, and dialogic management of classroom disagreements. This bifactorial structure, grounded in the cognitive-ethical and socio-political dimensions of argumentative competence, underpins hypotheses H1 and H2, which aim to confirm a two-dimensional structure through CFA.

The hypotheses formulated pursue three main objectives: first, to confirm the underlying/factorial structure of the latent factors through confirmatory factor analysis (CFA) ( $H_1$  and  $H_2$ ); second, to examine the potential statistical association of the predictors (teachers' geographic origin, age, and gender) with the latent factors of the AASES instrument (Assessment of Argumentation in Social and Experimental Sciences) ( $H_3$ ); and third, to explore the role of gender as a statistical mediating variable in the relationship between age and geographical origin with the scores obtained in the latent factors ( $H_4$  and  $H_5$ ). Figure 1 follows standard SEM graphic conventions: rectangles for manifest indicators, ovals or circles for latent constructs, and

arrows representing hypothesized causal or correlational relationships (Hair et al. 2022, 2024).

In the present study, the variable gender refers exclusively to the participating teachers (self-identified as male=1 or female=2), consistent with its treatment as a nominal dichotomous variable; no data on students' gender were collected. The rationale for proposing this factor as a mediating variable lies in previous literature. Indeed, gender may function as a mediating variable by interacting with other factors to shape argumentative abilities. Tsemach and Zohar (2021) found that gender differences in the construction of arguments become more or less visible depending on cultural context and prior training. This mediating role can also be observed in the quality of collaboration and argumentative outcomes; although gender alone does not determine performance, it does moderate the relationship between pedagogical context and academic achievement (Hsu et al. 2017). Likewise, Banihashem et al. (2024) showed that certain epistemic beliefs enhance argumentation only when gender amplifies or attenuates their effect, establishing a bidirectional mediation: gender influences how these beliefs are adopted, and in turn, these beliefs shape how men and women argue. Similarly, Noroozi et al. (2023) found that receptiveness to peer review feedback varies by gender. This finding suggests that the adoption and application of critical feedback depend not only on prior experience or individual disposition, but also on specific cultural norms and beliefs about knowledge, which are differentially shaped by gender.

Additionally, the study examines intergroup differences based on gender, age, and geographical origin ( $H_6$ ), three factors commonly defined as independent variables in educational research due to their potential influence on teaching-learning experiences and their relevance in identifying inequalities or specific educational needs. Figure 2 expands upon this schema and displays the placement of each hypothesis within the model.

**H1:** Observable indicators (C1 to C5) load significantly onto Factor 1, confirming the existence of an underlying latent factor.

**H2:** Observable indicators (C6 to C10) load significantly onto Factor 2, confirming the existence of a second underlying latent factor.

**H3:** Gender, origin, and age are statistically associated with Factors 1 and 2.

**H4:** Age and origin are statistically associated with gender as observed in the sample.

**H5:** Gender acts as a statistical mediator between age, origin, and Factors 1 and 2.

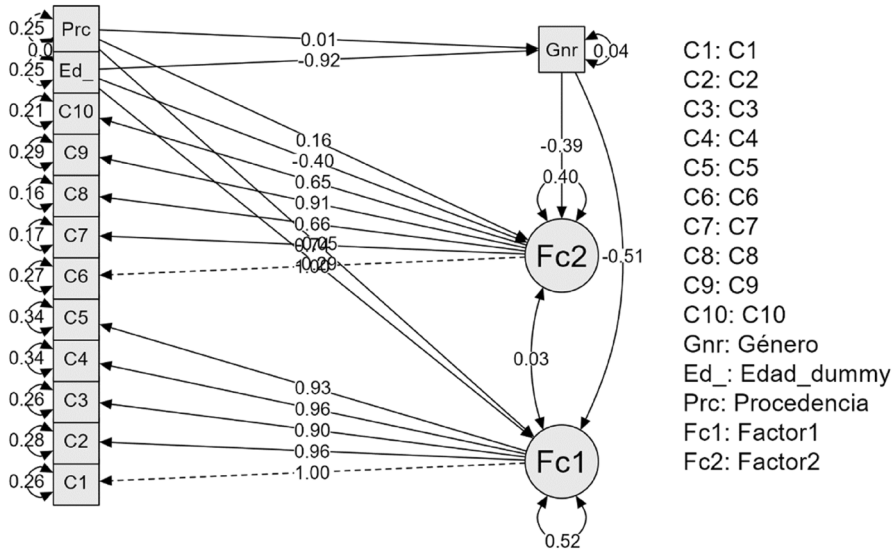
**H6:** There are statistically significant differences among the sociodemographic groups (gender, age, and geographical origin) in the perceptions of social science and science teachers regarding theoretical-factorial constructs 1 and 2.

## 2 Method

### 2.1 Participants

This study employed a non-probabilistic purposive sampling method. The sample size estimation for the optimal development of the factorial model was conducted

Proposed SEM model for the AASES instrument



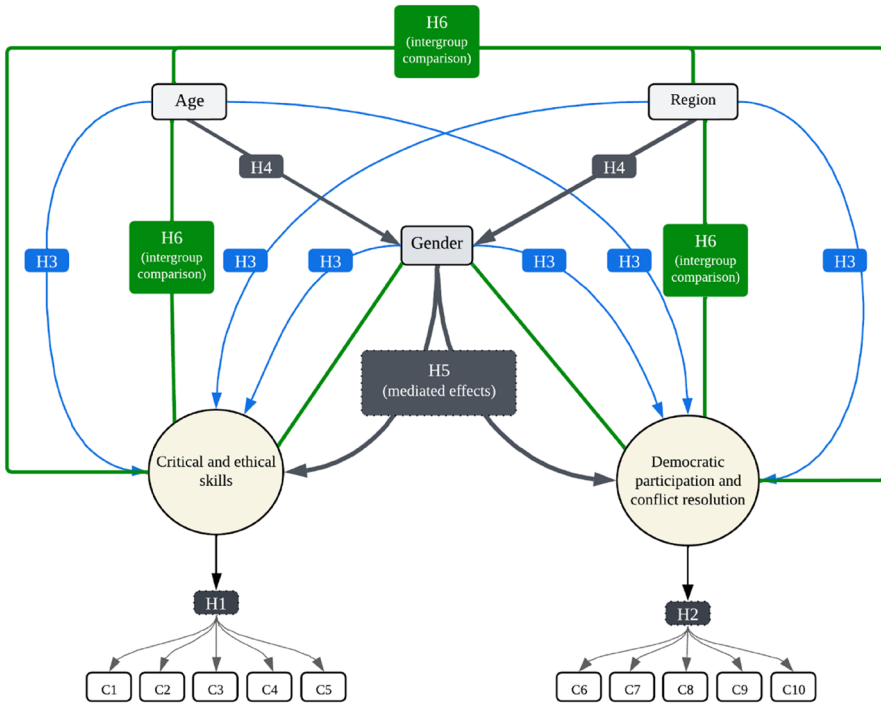
**Fig. 1** Proposed SEM model for the AASES instrument. *Note.* Squares represent observed variables (items Q01–Q10 and sociodemographic predictors: Gnr=gender, Ed\_ = dichotomized age, Prc=origin). Circles indicate latent variables (Fc1 = Critical and Ethical Skills; Fc2=Democratic Participation and Conflict Resolution). Solid unidirectional arrows represent regression paths (direct effects), and curved bidirectional arrows indicate correlations between latent factors. The numbers on each path correspond to standardized coefficients

using Structural Equation Modeling (SEM), through Soper’s advanced online calculator (Ventura-León et al. 2020). Based on 10 observed variables, 2 latent variables, an expected effect size of  $\lambda=0.3$  (the minimum acceptable  $\lambda$  value for factorial models), a desired probability level of 0.05, and a desired statistical power of  $1-\beta=0.95$ , the minimum recommended sample size was 147 cases. Taking this requirement into account, a total of 215 social science and science teachers from the eastern and southern regions of Spain participated in the study (Table 1).

### 2.2 Instrument

The AASES (Assessment of Argumentation in Social and Experimental Sciences) was used. This instrument, specifically designed for the study, consists of 10 items measured on a 7-point Likert scale (1=Strongly Disagree, 7=Strongly Agree). It is aimed at assessing the perceptions of in-service teachers in social and experimental sciences regarding argumentative competence, and is structured into two theoretical subconstructs (bifactor model):

Factor 1 (C1–C5): Critical and Ethical Skills, focused on measuring the perceived importance of argumentation in the development and acquisition of critical thinking skills, informed decision-making, and ethical discussion within the context of teaching social and experimental sciences (e.g., “Argumentation in the integrated teaching



**Fig. 2** Path model of the proposed hypotheses. *Note.* The yellow ovals represent the model's latent factors: Factor 1 – Critical and Ethical Skills and Factor 2 – Democratic Participation and Conflict Resolution

Source: Own elaboration.

of social and experimental sciences has a substantial impact on students' ability to make more informed decisions on social, socio-scientific, and political issues”).

Factor 2 (C6–C10): Democratic Participation and Conflict Resolution, containing items designed to assess the role of argumentation in fostering active participation in democratic life, conflict resolution, and the need to adapt arguments to different audiences and contexts (e.g., “The teaching of argumentation should focus on the development of students’ skills to adapt their arguments to different audiences and contexts”) (Table 2).

### 2.2.1 Content Validity

The quality of the item content was assessed through the participation of six external expert judges, intentionally selected based on their specialization in the fields of Social Science Education ( $n=2$ ), Science Education ( $n=2$ ), and General Linguistics ( $n=2$ )—three areas of knowledge closely related to the proposed research problem. These judges evaluated, on a 4-point scale, the following criteria: *relevance* (the conceptual alignment of the item with the domain intended to be assessed), *representativeness* (the extent to which the item content reflects the full range of characteristics of the construct), *clarity* (linguistic clarity, appropriate vocabulary use,

absence of ambiguity, and general comprehensibility of the item), *cultural and contextual relevance* (the suitability of the item for the target population, considering its educational, cultural, and linguistic background), and *absence of irrelevant or biased content* (lack of elements not directly related to the construct and avoidance of wording biases that may lead responses) (Pérez 2011; Valdés et al. 2019).

Based on the results, Bangdiwala's weighted concordance indices ( $B^W_N$ ) were calculated. These indices not only quantify the strength of agreement but also offer a graphical representation (Fig. 3). In this representation, black squares indicate observed agreement, grey squares indicate partial agreement, and the white area within each rectangle represents disagreement. According to Bangdiwala and Shankar (2013),  $B^W_N$  values are interpreted as follows: 0.00–0.20 indicates poor agreement; 0.20–0.40, weak; 0.40–0.60, moderate; 0.60–0.80, good; and values above 0.80, excellent.

The overall values obtained were satisfactory, both in the intra-group comparisons among judges from the same area of expertise and in the random inter-group combinations ( $B^W_N \geq 0.93$ ). Likewise, Gwet's AC2 coefficient was used to evaluate inter-rater consistency in the global assessment of the five content validity criteria. AC2 values reached  $\geq 0.93$  ( $p < .001$ ), indicating very high agreement among judges beyond chance levels (Table 3).

## 2.2.2 Internal Consistency Validity

The internal consistency reliability of the instrument's items was estimated using Cronbach's alpha ( $\alpha$ ) and McDonald's omega ( $\omega$ ) coefficients. First, reliability was calculated for the overall scale, and subsequently, for each of the proposed theoretical factors or subconstructs. The results indicated acceptable values (Table 4).

## 2.3 Design and Procedure

This study employed a non-experimental, cross-sectional research design, suitable for exploring relationships and comparing groups without manipulating the variables of interest. Although this design does not allow for causal inference due to the absence of experimental manipulation and control, it provides a descriptive and meaningful perspective on the trends present in the population studied.

**Table 1** Sociodemographic characteristics

		$f_i (p_i)$
Gender	Male	103 (47.9%)
	Female	112 (52.1%)
	Other identities	0 (0%)
Region	Southern Spain	113 (52.6%)
	Eastern Spain	102 (47.4%)
Age	Age group 11 ( $\geq 25 - \leq 35$ )	54 (25.1%)
	Age group 2 ( $\geq 36 - \leq 46$ )	49 (22.8%)
	Age group 3 ( $\geq 47 - \leq 57$ )	55 (25.6%)
	Age group 4 ( $\geq 58$ )	57 (26.5%)

**Table 2** AAES: items, factors, and operational criteria

Factor 1. Critical and Ethical Skills		
		Criterion C-x
Q01	Teaching argumentation contributes to the development of critical thinking within the specific context of integrated teaching of social and experimental sciences.	C1 Critical thinking
Q02	Teaching argumentation has a significant impact on students' critical capacity to communicate their ideas effectively.	C2 Critical communication
Q03	In most situations, students experience substantial benefits from acquiring critical inquiry skills to support their arguments in social and experimental sciences.	C3 Critical inquiry
Q04	Argumentation in the integrated teaching of social and experimental sciences has a substantial impact on students' ability to make more informed decisions on social, socio-scientific, and political issues.	C4 Informed decision-making
Q05	Learning argumentation has a considerable impact on fostering the importance of ethical discussion in controversial social and socio-scientific topics.	C5 Ethical discussion
Factor 2. Democratic Participation and Conflict Resolution		
Q06	Integrated teaching of social and experimental sciences effectively prepares students to use argumentation for active participation in democratic life.	C6 Democratic participation
Q07	The teaching of argumentation should focus on developing students' skills to adapt their arguments to different audiences and contexts.	C7 Argument adaptation
Q08	Argumentation in the integrated teaching of social and experimental sciences tends to foster a deeper understanding of the social concepts and theories underlying controversial socio-scientific topics.	C8 Deep understanding of socio-scientific concepts
Q09	Argumentation plays a fundamental role in the constructive resolution of conflicts and disagreements in the social sphere.	C9 Conflict resolution
Q10	Argumentation teaching should be given special consideration in the cross-curricular citizenship education program in secondary education.	C10 Curricular inclusion

The instrument was administered electronically to the participant sample between September 2024 and January 2025, following prior communication regarding the study's objectives and the confidentiality of the data collected. The questionnaire was distributed via the Google Forms platform, through a link sent by institutional email to the departments of Social and Experimental Sciences. Informed consent was also requested to authorize the use of participants' responses. Participation was entirely anonymous and voluntary, and no financial or academic compensation was offered, in order to ensure neutrality and avoid incentives that could bias the responses.

The study was conducted in accordance with the principles outlined in the Declaration of Helsinki (World Medical Association Declaration), ensuring a strong ethical-philosophical commitment and unwavering respect for human dignity, privacy, physical and moral integrity, and the protection of personal data throughout the survey and the research process. Moreover, the project was reviewed and approved by the Bioethics Committee of the university responsible for the research (RI 15/2018).

## 2.4 Data Analysis

A Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM) were conducted to validate the underlying empirical structure of the observed variables. Mediation analysis was performed to explore indirect relationships among variables and to propose an explanatory model. In addition, comparative-relational

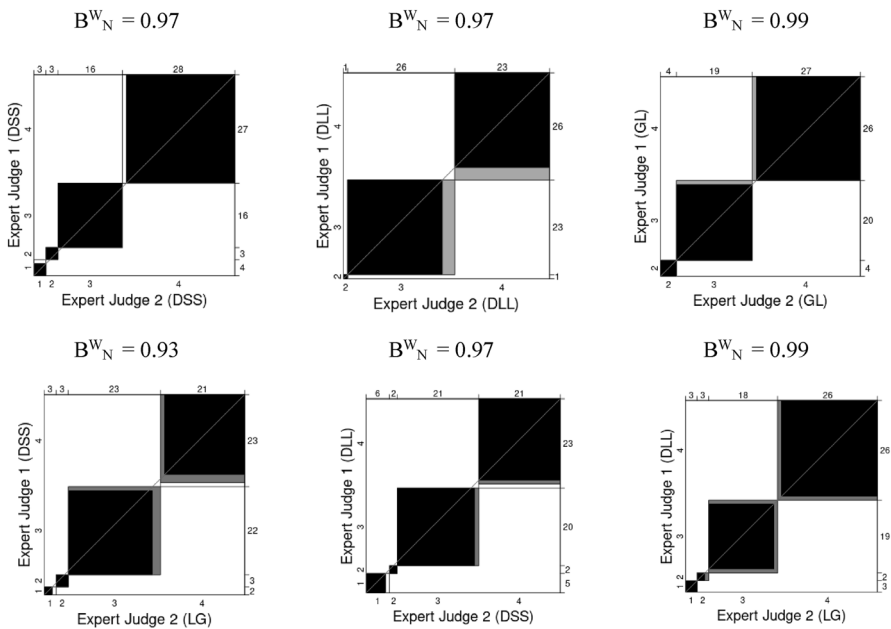


Fig. 3 Expert judge agreement coefficient ( $B^W_N$ )

Table 3 Expert judge agreement coefficient (Gwet's AC2)

EJ: Expert Judge; SSE: Social Science Education; SE: Science Education; GL: General Linguistics

EJ	$M_o$	Gwet's AC2	95% CI		$p$
			Lower	Upper	
SSE	4	0.96	0.869	1	<0.001
SE	3	0.96	0.916	1	<0.001
GL	4	0.99	0.957	1	<0.001
SSE-GL	3	0.93	0.84	1	<0.001
SE-SSE	4	0.95	0.849	1	<0.001
SE-LG	4	0.98	0.953	1	<0.001

Table 4 Internal consistency coefficients

AASES		Critical and ethical skills		Democratic participation and conflict resolution	
$\omega$	$\alpha$	$\omega$	$\alpha$	$\omega$	$\alpha$
0.804	0.796	0.891	0.890	0.855	0.849

inferential analyses were carried out to identify statistically significant differences between various sociodemographic groups in relation to the study variable.

The Kolmogorov–Smirnov test was significant ( $p < .05$ ), indicating a rejection of strict normality. However, the skewness coefficients for each factor and comparison group were below 2 ( $S = -1.19$  to  $-0.291$ ,  $SE = 0.166$  to  $0.166$ ), and the kurtosis coefficients were below 7 ( $S = -0.699$  to  $0.301$ ,  $SE = 0.330$  to  $0.330$ ), suggesting that the distribution approximates normality. Based on this descriptive evidence of normal-

ity and the verification of homoscedasticity across group distributions, parametric procedures were employed for group comparisons (one-way ANOVAs), using the following dependent variables the theoretical dimensions of the administered instrument, and the sociodemographic groups of the participant sample as the independent variable.

Additionally, effect size ( $f$ ) was calculated to determine the strength and magnitude of the differences identified, as well as statistical power, to minimize the risk of Type II errors. Finally, Bonferroni's *post-hoc* multiple comparison test was applied to determine the specific location of the expected intergroup differences.

### 3 Results

#### 3.1 Confirmatory Model Evaluation: Testing Hypotheses 1 and 2 Through Fit Indices

The fit indices obtained indicate that the model demonstrates an excellent fit. The Comparative Fit Index (CFI=0.98) and the Tucker–Lewis Index (TLI=0.99) exceed the recommended threshold of 0.95. The Incremental Fit Index (IFI) and the Relative Noncentrality Index (RNI) also return values of 0.98, surpassing the 0.95 benchmark. The Root Mean Square Error of Approximation (RMSEA) is 0.00, with a confidence interval of 0.00, indicating a total absence of discrepancy between the model and the data. The RMSEA was 0.00 (CI=0.00). Although values below 0.05 indicate a good fit, an exact value of 0.00 is statistically uncommon in real data and may be attributed to the high parsimony of the model and the specific characteristics of the sample. Therefore, it should be interpreted as evidence of excellent fit, without assuming a perfect fit (Valdés et al. 2019).

In addition, the Standardized Root Mean Square Residual (SRMR) is 0.009, well below the conventional cut-off of 0.08. Lastly, the Goodness of Fit Index (GFI) is 1, and the Adjusted Goodness of Fit Index (AGFI) is 0.996, both close to or equal to 1, further supporting the quality of model fit. Given that these indices reflect the model's alignment with empirical data, the proposed theoretical structure—with two dimensions (Factor 1 [C1–C5]: *Critical and Ethical Skills*, and Factor 2 [C6–C10]: *Democratic Participation and Conflict Resolution*)—can be confirmed (Fig. 4).

Despite the general acceptability of the error variances ( $<0.30$ ), they indicate that less than approximately 80% of the item variance is explained by the corresponding factor. This suggests room for improvement in measurement quality or that the item may not be fully aligned with the theoretical construct it is intended to represent. In contrast, the observed factor loadings are satisfactory (approximately  $>0.50$ ), indicating that the items represent their respective factors well. Finally, a very weak correlation between the latent factors was observed ( $r=.07$ ), confirming their independence.

### 3.2 Structural Regression Model Evaluation: Testing Hypotheses 3 to 5 on the Effect of Sociodemographic Variables

It was found that age, after being dichotomized into a dummy variable, shows a statistically significant association with gender ( $\beta = -0.920, p < .001; 95\% \text{ CI } [-0.972, -0.867]$ ). The confidence interval  $(-0.972, -0.867)$  does not include zero, confirming the significance of the negative effect. This indicates that, in this sample, the older age group (over 58 years) was overrepresented by male participants.

These results suggest that age is the only variable with a significant effect on gender—a necessary condition for mediation. However, it can also be observed that gender has no significant effect on the latent factors ( $p > .05$ ), meaning its mediating role in the relationship between age, region, and the latent factors cannot be confirmed. Since none of the predictors (age, region, gender) significantly explained the latent factors, their predictive functions must be ruled out, as they do not contribute meaningfully to the explained variance. Consequently, none of these predictors appears to be useful in accounting for the proposed latent constructs.

### 3.3 Analysis of Variance: Testing Hypothesis 6 on Differences in Latent Factors by Gender, Age, and Region

The results reveal statistically significant differences in the *Critical and Ethical Skills* dimension (Factor 1) based on the gender identities expressed by participants ( $F_{(1, 213)} = 5.49, p = .020, f = 0.15, 1 - \beta = 0.63$ ), with higher scores observed among male participants compared to female participants (Table 6).

These findings suggest a slightly more positive tendency among men to emphasize the importance of teaching argumentation for informed decision-making regarding controversial social and socio-scientific issues, ethical discussion of social matters,

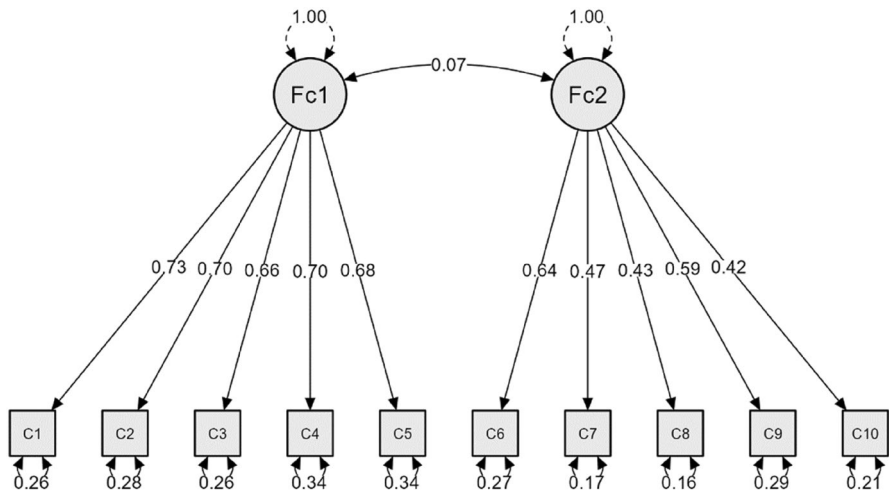


Fig. 4 Path diagram illustrating the confirmatory relationship between latent variables (factors) and their observed indicators

**Table 5** Regression coefficients

Predictor	Outcome	Estimate	SE	z	p	95% IC	
						Lower	Upper
Age	Factor1	-0.285	0.265	-1.074	0.283	-0.805	0.235
Region	Factor1	-0.050	0.104	-0.478	0.633	-0.254	0.154
Gender	Factor1	-0.505	0.266	-1.898	0.058	-1.027	0.016
Age	Factor2	-0.405	0.239	-1.690	0.091	-0.874	0.065
Region	Factor2	0.157	0.094	1.667	0.096	-0.028	0.341
Gender	Factor2	-0.388	0.239	-1.619	0.105	-0.857	0.082
Age	Gender	-0.920	0.027	-34.353	<0.001	-0.972	-0.867
Region	Gender	0.012	0.027	0.460	0.645	-0.040	0.065

Note. The statistical path age  $\rightarrow$  gender was included to meet the predictor  $\rightarrow$  mediator condition (estimation of the predictor  $\rightarrow$  mediator path) required by the mediation analysis. Its implications, therefore, do not suggest any potential determination of age over gender, but rather that a statistical association exists between these two variables within this sample. Thus, gender is identified as a mediating variable solely for modeling purposes; its position as an outcome does not imply any causal relationship with age or origin

the development of critical thinking, the improvement of effective communication, and the acquisition of critical inquiry skills. However, this difference reflects a small effect size, indicating a limited practical impact of the independent variable on the dependent one. Moreover, the statistical power did not reach the recommended threshold ( $<0.80$ ), which implies a considerable likelihood of failing to detect a real difference if one does exist, and a heightened risk of committing a Type II error.

Statistically significant differences were also found for the age variable in the same dimension (*Critical and Ethical Skills*, Factor 1 ( $F_{(3,211)}=4.59, p=.004, f=0.24, 1-\beta=0.87$ ) (Table 7), with notably higher scores in the oldest age group. Specifically, participants over 58 years of age rated the importance of argumentation teaching more positively regarding the variables mentioned. These differences achieved optimal statistical power values ( $\geq 0.80$ ) and medium effect sizes ( $f \geq 0.24$ ).

Intergroup differences were observed between the 36–46 and 58+ age groups ( $p=.013$ , 95% CI  $[-0.809, -0.062]$ ), and between the 47–57 and 58+ age groups ( $p=.009$ , 95% CI  $[0.073, 0.798]$ ). However, no statistically significant differences were found between gender ( $p=.84$ ) and age ( $p=.14$ ) subgroups in relation to the dimension *Democratic Participation and Conflict Resolution* (Factor 2), which is associated with the capacity of argumentation teaching to foster active democratic engagement, adapt arguments to different audiences and contexts, comprehend social theories, resolve conflicts constructively, and promote essential ethical and critical citizenship skills. Likewise, the levels of the region variable did not reveal statistically significant differences between the two demographic groups ( $F_1, p=.624$ ;  $F_2, p=.123$ ). Descriptive mean values were similar for teachers from southern Spain ( $n=113$ ;  $F_1: M=6.03, SD=0.79$ ;  $F_2: M=6.20, SD=0.57$ ) and from eastern Spain ( $n=102$ ;  $F_1: M=5.98, SD=0.67$ ;  $F_2: M=6.32, SD=0.52$ ).

**Table 6** Descriptive and Inferential statistics for the gender variable

	Gender		F	gl <sub>1</sub>	gl <sub>2</sub>	p	f	1-β
	♂ (n=103)	♀ (n=112)						
	M (DT)	M (DT)						
F1	<b>6.13 (0.70)</b>	5.90 (0.75)	5.49	1	213	<b>0.020*</b>	0.15	0.63
F2	6.27 (0.55)	6.26 (0.55)	-	-	-	-	-	-

\* $p < .05$ 

## 4 Discussion and Conclusions

This study identified two main descriptive factors regarding the importance of argumentation in the integrated teaching of social and experimental sciences: one related to critical and ethical skills, and the other associated with democratic participation, engagement with socio-scientific issues, and conflict resolution. These findings suggest, according to the perceptions of the participating teachers, that argumentation may contribute to the development of reflective and civically engaged students with their natural, social, and cultural environment (Bernal-Cavieres and Villalón 2024), and supports the pedagogical integration of disciplinary content with the development and acquisition of deliberative skills (Rapanta and Macagno 2016). However, given the cross-sectional nature of the study and its reliance on self-reported data, it is not possible to establish causal effects on students; the data only reflect teachers' beliefs. The first factor emphasizes the relevance of argumentation in promoting critical thinking, the responsible use of evidence, and ethical understanding of socio-scientific problems, in line with Macagno's (2016) thesis, which highlights both the structure and discursive relevance of arguments. The second factor refers to active participation in democratic life and the capacity to manage disagreements constructively, consistent with research that underscores the dialogic and cooperative nature of argumentation in decision-making (Kuhn and Halpern 2022) and in facilitating discussions on complex issues (Asterhan and Schwarz 2016). The theoretical constructs revealed in the AASES scale thus anticipate the fundamental tension between epistemic accuracy and deliberative openness described by Zenker et al. (2024), conceptualized as the ideals of correctness and participation in public argumentation. This parallel confirms that what educators value (the formation of critical thinkers with deliberative competencies) essentially corresponds to the curricular translation of the two normative pillars that structure contemporary civic deliberation.

This research also highlights the urgent need to establish a curriculum capable of incorporating argumentation in a transversal manner, as well as the importance of teacher training in implementing instructional strategies aimed at fostering deliberation and critical analysis in the classroom (Rapanta and Felton 2022). To this end, the study provides specific data on how *age* and *gender* moderately influence the evaluation of the critical-ethical dimension of argumentation teaching—an aspect that has received comparatively less attention in the international literature, which has largely focused on factors such as pedagogical approaches, instructional planning, the quality of teacher guidance, and learning environments that encourage reflection and collaborative dialogue (Iordanou and Constantinou 2015).

**Table 7** Descriptive and Inferential statistics for the age variable

Age group	2 ( $n = 49$ )		3 ( $n = 55$ )		4 ( $n = 57$ )		F	gl <sub>1</sub>	gl <sub>2</sub>	$p$	$f$	$1-\beta$	$f_1$	$1-\beta_1$	$f_2$	$1-\beta_2$
	$M(DT)$	$M(DT)$	$M(DT)$	$M(DT)$	$M(DT)$	$M(DT)$										
F1	5.96 (0.68)	5.86 (0.83)	5.86 (0.70)	<b>6.30 (0.66)</b>	4.59	3	211	-	-	<b>0.004**</b>	0.24	0.87	0.28	0.82	0.30	0.88
F2	6.37 (0.48)	6.17 (0.58)	6.17 (0.54)	6.31 (0.57)	-	-	-	-	-	-	-	-	-	-	-	-

Note. 1 = 25–35 years; 2 = 36–46 years; 3 = 47–57 years; 4 = over 58 years

\*\* $p < .01$ .  $f_1$  = effect size between age groups 2 and 4.  $f_2$  = effect size between age groups 3 and 4

Indeed, teachers' perceptions of the educational value of argumentation were very positive, with mean scores approaching the upper end of the scale. However, the differences identified nuance this general consensus. Unlike the results of previous studies, in which women report higher levels of self-efficacy in teaching thinking skills (Dilekli and Tezci 2019), male teachers in this study reported slightly higher scores in the critical and ethical skills dimension, indicating a higher self-perception of competence in this area. In this regard, three possible explanations may account for the divergence from previous studies: (1) in the sample analyzed in this study, male teachers are overrepresented in the oldest age group ( $\geq 58$  years), which also obtained the highest scores in *Factor 1 – Critical and Ethical Skills*; this age-gender overlap may be inflating the male average, even though the effect size is small ( $f=0.15$ ) and statistical power is moderate ( $1-\beta=0.63$ ); (2) the use of self-report measures may result in potential social desirability biases and, consequently, in a possible overestimation of competence, which the literature has linked to gender-based differences,—particularly in the assessment of teacher self-efficacy (Benito et al. 2010); (3) according to the post hoc test results, the difference is statistically significant but of limited practical impact, suggesting that sample variability may also be influencing the outcome. These considerations indicate that the finding should be interpreted with caution and verified in future studies using larger and more balanced samples.

Moreover, older teachers,—particularly those over the age of 58,—reported significantly higher scores in this same dimension, possibly due to the accumulation of professional experience, which may foster greater awareness of the positive impact of argumentation on the development of critical thinking, informed decision-making, and ethical discussion. Within this age group, it is male teachers who report the highest scores in *Factor 1 – Critical and Ethical Skills*, suggesting a possible combined effect of gender and professional seniority on the perceptions analyzed. Along similar lines, Ulu-Kalın and Kumandaş-Öztürk (2024) found that teachers with greater professional seniority scored significantly higher in creative thinking (problem-solving). This conclusion suggests, therefore, that the competence profile evolves over time and in response to the demands of the educational context. These differences highlight that the valuation of argumentation is not homogeneous but is instead modulated by factors such as gender and professional seniority, which reinforces the need for both initial and ongoing teacher training that is sensitive to these variables.

The absence of statistically significant differences in relation to geographical origin regarding the willingness to employ specific argumentation techniques in the classroom aligns with the notion that the effectiveness of argumentative practices is more directly dependent on institutional support and teachers' willingness to incorporate structured debate dynamics (Pérez-Echeverría et al. 2016). Alongside the need to design and administer instruments with sufficient empirical guarantees of validity and reliability—offered here for the first time in the present study—it is also essential to further investigate the sociodemographic variables that characterize the teaching population, in order to better understand the differential aspects preliminarily interpreted in this analysis.

This study also considered the implementation of argumentation within the integrated field of Social and Science Education, and argumentative writing around scientific controversies (Iordanou and Constantinou 2015; Casado-Ledesma et al.

2021), reinforcing the need to connect didactic knowledge with deliberative practice and transdisciplinary metacognitive reflection. Both the findings of this study and previous evidence concur in affirming that argumentation is an essential competence for educating critically engaged and participatory citizens, and that its development requires continuous teacher training focused on the design, implementation, and evaluation of specific instructional resources.

In this regard, the results obtained suggest the need to implement teacher training programs or modules with an explicit focus on argumentation. From this perspective, curricula at various educational levels and teacher preparation programs should offer appropriate and more targeted training spaces for the development and acquisition of skills and competencies related to argumentation and informed engagement with social issues, current social debates, or contemporary socio-scientific controversies.

Argumentation represents a central component in the development of critical thinking skills and in the education of active citizenship. It enhances critical thinking, fosters the co-construction of meaning, and supports emotional regulation within social and civic contexts. Teaching argumentation, therefore, entails selecting the most appropriate theoretical models for each educational level and discipline, designing environments that promote authentic dialogue, providing explicit instructional support, encouraging reflection on one's own argumentative process, and employing diverse forms of assessment that capture the complexity of reasoning.

Teaching and learning to argue thus enhances the quality of written texts and generated debates, incorporates metacognition as a key mediating factor between argumentative practice and meaningful learning, and contributes to shaping more critical and reflective identities across disciplinary domains. Accordingly, the importance of argumentation—as a transversal, transdisciplinary competence—requires the integration of instructional strategies grounded in socially situated practice and the articulation of adaptive scaffolding aimed at promoting reflection and the co-construction of meaning.

The educational relevance of argumentation and the transdisciplinary functionality of deliberative activities for fostering critical and contextualized discussions,—even in highly polarized social contexts (Bernal-Cavieres and Villalón 2024),—underscore the need for mutual reinforcement between argumentation theory and educational practice, as well as for a deeper exploration of pedagogical strategies that promote structured dialogue, reflection, and evidence-based reasoning. These results entail direct curricular implications. The high value that teachers attribute to argumentation, along with the variations observed by age and gender, reveal that, in the absence of precise curricular guidelines, the presence of argumentative practices is largely dependent on the individual teacher's initiative and training. Therefore, it is essential to incorporate specific objectives, content, and assessment criteria on argumentation into the Secondary Education decrees, in alignment with the development and acquisition of the key competencies outlined in the Spanish LOMLOE (Organic Law amending the Organic Law of Education) and with the inherent democratic dimension of citizenship education. In this way, the curriculum would serve as a stable and legitimizing framework for argumentative practices and as a guiding reference for the initial, continuous, and ongoing training of teachers. A true cross-curricular integration of argumentation is therefore essential, along with increased teacher training

focused on designing activities and scaffolds that enable students to understand, use, and refute evidence constructively.

#### 4.1 Limitations and Future Research Directions

One of the main limitations of the present study lies in its non-experimental, cross-sectional design. This methodological approach limits the ability to understand how teachers' perceptions of argumentation evolve over time or across different contexts. Furthermore, although the mediating role of gender between age, geographical origin, and the latent factors defined in the instrument was analyzed, the results do not support the confirmation of such mediation. This absence of mediating effects may stem either from limitations in the theoretical model itself or from an insufficiently sensitive operationalization of the variables involved.

The statistical analysis also revealed that the sociodemographic predictors (gender, age, and geographical origin) do not significantly explain the latent factors measured by the AASES instrument. This lack of association may reflect limitations in the design of the instrument itself, as well as the need to incorporate additional contextual or psychoeducational variables that offer greater explanatory power. It is also important to note the restricted generalizability of the results, given that the sample was limited to secondary school teachers in Spain and did not include other educational levels or international contexts.

Considering these limitations, future research should explore the use of longitudinal designs aimed at observing potential changes in teachers' perceptions over time, taking into account their professional experience or the implementation of targeted teaching interventions. In addition, to gain a deeper understanding of the contextual factors influencing perceptions of argumentation, it would be valuable to conduct complementary qualitative studies—such as in-depth interviews or focus groups—to uncover the subjective reasons underlying the differences observed by gender or age. In this regard, qualitative research could help identify meanings and beliefs that are not always accessible through structured questionnaires (Sáiz Manzanares et al. 2019).

Another promising avenue for future research involves conducting cross-cultural comparative studies to examine how these perceptions are shaped in different educational systems, which could offer culturally grounded insights for interpreting the results. From a methodological perspective, the use of multigroup structural equation modeling could allow for more precise comparisons between sociodemographic segments, while also incorporating latent variables such as initial training in citizenship education or teaching experience in argumentative instruction. Likewise, it would be especially valuable to incorporate students' perspectives, either by adapting the AASES instrument or through the use of qualitative techniques, in order to assess the extent to which their evaluations converge with or differ from those of their teachers. This would contribute to a more holistic and comprehensive understanding of argumentation-related teaching and learning processes in the classroom.

Finally, it is recommended to promote transdisciplinary research projects in which argumentation,—understood as a key competence for critical citizenship,—is approached from multiple perspectives. This would enable the design of robust train-

ing programs that are sensitive to diverse cultural and social contexts. In this line, the design and implementation of educational intervention programs aimed at improving teachers' competences for teaching argumentative skills should be encouraged, with their impact assessed through the AASES instrument itself.

**Funding** Open access funding provided by FEDER European Funds and the Junta de Castilla y León under the Research and Innovation Strategy for Smart Specialization (RIS3) of Castilla y León 2021-2027. This research was funded by the project Attitudes, didactic predisposition, and interests of prospective teachers regarding the curricular inclusion of controversial issues in compulsory secondary education (ADIFICE) (PID2024-156838OB-I00), supported by the State Research Agency of the Spanish Ministry of Science, Innovation and Universities; by the project The Nature of I-STEM (NoSTEM) for Citizen Training (PID2020-118010RB-I00), supported by the State Research Agency of the Spanish Ministry of Science and Innovation; by the Consolidated Research Unit 285 – Integrated Teaching and Transdisciplinarity (EIT); and by the Recognized Research Group Didactics of History and Social Sciences (DHISO) at the University of Burgos.

**Data Availability** The datasets generated and/or analyzed during the current study are not publicly available due to privacy and confidentiality agreements with the participants. However, they may be made available by the corresponding author upon reasonable request, provided that the purpose is justified and appropriate measures are taken to safeguard the privacy and anonymity of the individuals involved.

## Declarations

**Conflict of interest** The authors have not reported any potential conflicts of interest in relation to this article.

**Ethical Approval and Informed Consent** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards, and also with the standards of research ethical criteria of the University of Burgos (Spain). The Bioethics Commission of the University of Burgos approved the research project on June 21, 2024 thus guaranteeing the ethical-philosophical commitment and indeclinable respect for human dignity, privacy, physical and moral integrity as well as the protection of personal data in the treatment of the survey and throughout the course of the research. Informed participation is guaranteed, as well as publication consent. All participants were adults and their personal data has been anonymized.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Altun, E., and T. Ozsevgec. 2025. Making argumentation-based learning and teaching happen: Exploring the development of pre-service science teachers' argumentation competencies. *Science & Education*. <https://doi.org/10.1007/s11191-024-00612-1>.

- Asterhan, C. S. C., and B. B. Schwarz. 2016. Argumentation for learning: Well-trodden paths and unexplored territories. *Educational Psychologist* 51(2):164–187. <https://doi.org/10.1080/00461520.2016.1155458>
- Bangdiwala, S. I., and V. Shankar. 2013. The agreement chart. *BMC Medical Research Methodology* 13:97. <https://doi.org/10.1186/1471-2288-13-97>
- Banihashem, S. K., O. Noroozi, H. J. A. Biemans, and V. C. Tassone. 2024. The intersection of epistemic beliefs and gender in argumentation performance. *Innovations in Education and Teaching International* 61(4):716–734. <https://doi.org/10.1080/14703297.2023.2198995>
- Benito, J. G., M. D. H. Montesinos, and G. Guilera. 2010. El Sesgo de Los instrumentos de medición: Tests Justos. *Papeles Del Psicólogo* 31(1):75–84.
- Bernal-Cavieles, E., and G. Villalón. 2024. Profesores, deliberación y Contenidos históricos En contexto de Estallido social. *Educación Y Humanismo* 26(46):181–198. <https://doi.org/10.17081/eduhum.26.46.6630>
- Blair, J. A. 2023. Teaching the fallacies. *Argumentation* 37(2):247–251. <https://doi.org/10.1007/s10503-023-09604-x>
- Bubikova-Moan, J., and M. Sandvik. 2022. Argumentation in early childhood: A systematic review. *Human Development* 66(6):397–413. <https://doi.org/10.1159/000527293>
- Canals, R. 2007. La Argumentación En El Aprendizaje Del Conocimiento social. *Enseñanza De Las Ciencias Sociales: Revista De Investigación* 6:49–60. <https://raco.cat/index.php/EnsenanzaCS/article/view/126328>
- Casado-Ledesma, L., I. Cuevas, and E. Martín. 2021. Learning science through argumentative synthesis writing and deliberative dialogues: A comprehensive and effective methodology in secondary education. *Reading and Writing* 36(4):965–996. <https://doi.org/10.1007/s11145-021-10191-0>
- Dilekli, Y., and E. Tezci. 2019. A cross-cultural study: Teachers' self-efficacy beliefs for teaching thinking skills. *Thinking Skills and Creativity* 34:100624. <https://doi.org/10.1016/j.tsc.2019.100624>
- Evagorou, M., M. Vrikki, and E. Papanastasiou. 2023. Students' and teachers' voice on the outcomes of a citizenship education curriculum. *Citizenship, Social and Economics Education* 22: 100–117. <https://doi.org/10.1177/14788047231193917>
- Fuentes-Moreno, C., M. Sabariego-Puig, and A. Ambrós-Pallarés. 2020. Developing social and civic competence in secondary education through the implementation and evaluation of teaching units and educational environments. *Humanities and Social Sciences Communications* 7. <https://doi.org/10.1057/s41599-020-0530-4>
- Hair, J. F., G. T. M. Hult, C. M. Ringle, and M. Sarstedt. 2022. *A primer on partial least squares structural equation modeling (PLS-SEM)*. 3rd ed. Sage.
- Hair, J. F., M. Sarstedt, C. M. Ringle, and S. P. Gudergan. 2024. *Advanced issues in partial least squares structural equation modeling (PLS-SEM)*. 2nd ed. Sage.
- Hsu, P.-S., M. Van Dyke, and T. J. Smith. 2017. The effect of varied gender groupings on argumentation skills among middle school students in different cultures. *Middle Grades Review* 3(2):4. <http://scholarworks.uvm.edu/mgreview/vol3/iss2/4>
- Iordanou, K. 2022. Supporting strategic and meta-strategic development of argument skill: The role of reflection. *Metacognition and Learning* 17:399–425. <https://doi.org/10.1007/s11409-021-09289-1>
- Iordanou, K., and C. P. Constantinou. 2015. Supporting use of evidence in argumentation through practice in argumentation and reflection in the context of SOCRATES learning environment. *Science Education* 99(2):282–311. <https://doi.org/10.1002/sc.21152>
- Iordanou, K., and C. Rapanta. 2021. Argue with me: A method for developing argument skills. *Frontiers in Psychology* 12: 631203. <https://doi.org/10.3389/fpsyg.2021.631203>
- Kuhn, D., and M. Halpern. 2022. How might argumentation research inform discourse-based social studies education? *The Social Studies* 113(5):264–270. <https://doi.org/10.1080/00377996.2022.2053832>
- Kuhn, D., N. Zillmer, A. Crowell, and J. Zavala. 2013. Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentative competence. *Cognition and Instruction* 31(4):456–496. <https://doi.org/10.1080/07370008.2013.830618>
- Macagno, F. 2016. Argument relevance and structure: Assessing and developing students' uses of evidence. *International Journal of Educational Research* 79:180–194. <https://doi.org/10.1016/j.ijer.2016.07.002>
- Mallahi, O. 2024. Exploring the status of argumentative essay writing strategies and problems of Iranian EFL learners. *Asian-Pacific Journal of Second and Foreign Language Education* 9: 19. <https://doi.org/10.1186/s40862-023-00241-1>

- Martins, M., and F. Macagno. 2022. An analytical instrument for coding and assessing argumentative dialogues in science teaching contexts. *Science Education* 106(3):573–609. <https://doi.org/10.1002/sce.21708>
- Matos, F. 2021. Collaborative writing as a bridge from peer discourse to individual argumentative writing. *Reading and Writing* 34: 1321–1342. <https://doi.org/10.1007/s11145-020-10101-3>.
- Morawski, M., and A. Budke. 2019. How digital and oral peer feedback improves high school students' written argumentation—A case study exploring the effectiveness of peer feedback in geography. *Education Sciences* 9(3):178. <https://doi.org/10.3390/educsci9030178>
- Noroozi, O., S. K. Banihashem, N. T. Kerman, M. P. A. Khaneh, M. Babae, H. Ashrafi, and H. J. A. Biemans. 2023. Gender differences in students' argumentative essay writing, peer review performance and uptake in online learning environments. *Interactive Learning Environments* 31(10):6302–6316. <https://doi.org/10.1080/10494820.2022.2034887>
- Nussbaum, E. M. 2020. Critical integrative argumentation: Toward complexity in students' thinking. *Educational Psychologist* 56(1):1–17. <https://doi.org/10.1080/00461520.2020.1845173>
- Olmos, P. 2022. Argumentación y Pensamiento crítico: Convergencias y desafíos. *SCIO: Revista De Filosofía* 22:39–65. [https://doi.org/10.46583/scio\\_2022.22.992](https://doi.org/10.46583/scio_2022.22.992)
- Pérez, R. 2011. *Técnicas e instrumentos Para La recogida de Información*. Universidad Nacional de Educación a Distancia.
- Pérez-Echeverría, M.P., Y. Postigo, and M. García-Mila. 2016. Argumentation and education: Notes for a debate / Argumentación y educación: Apuntes Para Un debate. *Infancia y Aprendizaje* 39 (1): 1–24. <https://doi.org/10.1080/02103702.2015.1111607>.
- Rapanta, C. 2021. Can teachers implement a student-centered dialogical argumentation method across the curriculum? *Teaching and Teacher Education* 105:103404. <https://doi.org/10.1016/j.tate.2021.103404>
- Rapanta, C., and M. K. Felton. 2022. Learning to argue through dialogue: A review of instructional approaches. *Educational Psychology Review* 34:477–509. <https://doi.org/10.1007/s10648-021-09637-2>
- Rapanta, C., and F. Macagno. 2016. Argumentation methods in educational contexts: Introduction to the special issue. *International Journal of Educational Research* 79:142–149. <https://doi.org/10.1016/j.ijer.2016.03.006>
- Rapanta, C., M. García-Mila, and S. Gilabert. 2013. What is meant by argumentative competence?? An integrative review of methods of analysis and assessment in education. *Review Of Educational Research* 83 (4): 483–520. <https://doi.org/10.3102/0034654313487606>.
- Sáiz Manzanares, M., M. del Escobar Llamazares C., and J. Rodríguez Medina. 2019. *Investigación cualitativa. Aplicación de Métodos Mixtos y de técnicas de minería de Datos*. Universidad de Burgos.
- Sengul, O. 2019. Linking scientific literacy, scientific argumentation, and democratic citizenship. *Universal Journal of Educational Research* 7 (4): 1090–1098. <https://doi.org/10.13189/UJER.2019.070421>.
- Shi, Y. 2020. Constructed dialogs reveal skill development in argumentative writing. *Reading and Writing* 33:2311–2335. <https://doi.org/10.1007/s11145-020-10042-x>
- Spagnolo, C. 2024. Teachers' and students' beliefs on developing argumentative competence. *Edma 0–6: Educación Matemática en la Infancia* 13 (2): 60–72. <https://doi.org/10.24197/edmain.2.2024.60-72>.
- Tsemach, E., and A. Zohar. 2021. The intersection of gender and culture in argumentative writing. *International Journal of Science Education* 43(6):969–990. <https://doi.org/10.1080/09500693.2021.1894499>
- Ucar-Longford, B., A. Hosein, and M. Heron. 2024. Improving pre-service teachers' argumentation skills: A holistic online scaffolding design approach. *Journal of Digital Learning in Teacher Education* 40(2):76–96. <https://doi.org/10.1080/21532974.2024.2320440>
- Ulu-Kalin, Ö., and H. Kumandaş-Öztürk. 2024. Classification of higher-order thinking skills of the teachers based on institution, seniority, and branch with discriminant analysis. *Behavioral Sciences* 14 (8): 626. <https://doi.org/10.3390/bs14080626>.
- Valdés, A. A., F. I. García, G. M. Torres, M. Urías, and C. S. Grijalva. 2019. *Medición En investigación educativa Con Apoyo Del SPSS y ELAMOS*. Instituto Tecnológico de Sonora.
- Valero, A., O. Noroozi, H. Biemans, and M. Mulder. 2022. Argumentation competence: Students' argumentation knowledge, behavior and attitude and their relationships with domain-specific knowledge acquisition. *Journal of Constructivist Psychology* 35(1):123–145. <https://doi.org/10.1080/10720537.2020.1734995>
- van Eemeren, F.H., and B. Garssen. 2023. The pragmadialectical approach to the fallacies revisited. *Argumentation* 37 (2): 167–180. <https://doi.org/10.1007/s10503-023-09605-w>.

- Ventura-León, J. L., T. Caycho-Rodríguez, and M. Barboza-Palomino. 2020. ¿Cuántos participantes son necesarios? Un Método Para Estimar El Tamaño muestral En SEM. *Archivos Argentinos De Pediatría* 118(5):510.
- Vicente, J.J., A.J. Franco-Mariscal, and J.M. Oliva. 2024. Enhancing argumentation and decision-making of preservice early childhood education teachers through role-playing on animal experimentation. *Science & Education*. <https://doi.org/10.1007/s11191-024-00529-9>.
- Wang, L. J. 2025. Enhancing EFL argumentative writing: The impact of an online collaborative argumentation environment. *Computer Assisted Language Learning* 1–27. <https://doi.org/10.1080/09588221.2025.2482138>
- Zenker, F., J. A. van Laar, B. Cepollaro, A. Gâță, M. Hinton, C. G. King, B. Larson, M. Lewiński, C. Lumer, S. Oswald, M. Pichlak, B. D. Scott, M. Urbański, and J. H. M. Wagemans. 2024. Norms of public argumentation and the ideals of correctness and participation. *Argumentation* 38(1):7–40. <https://doi.org/10.1007/s10503-023-09598-6>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Delfin Ortega-Sánchez<sup>1</sup>  · Carlos Pérez-González<sup>2</sup> 

✉ Delfin Ortega-Sánchez  
dosanchez@ubu.es

Carlos Pérez-González  
cperez@ubu.es

<sup>1</sup> Department of Specific Didactics, Faculty of Education, University of Burgos, Burgos 09001, Spain

<sup>2</sup> Department of Philology, Faculty of Humanities and Communication, University of Burgos, Burgos 09001, Spain