

Article

Supporting Teachers' Understanding of Integrated STEM Through the Evaluation of a Teaching–Learning Sequence

V́ctor Mart́nez-Mart́nez , Jairo Ortiz-Revilla  and Ileana M. Greca 

Department of Specific Didactics, University of Burgos, 09004 Burgos, Spain; jortizr@ubu.es (J.O.-R.);
imgreca@ubu.es (I.M.G.)

* Correspondence: victormm@ubu.es

Abstract

The effective transfer of research-based instructional innovations into classroom practice requires assessment tools that allow teachers to critically examine the quality and applicability of STEM learning designs. This study employs the RubeSTEM rubric to analyze a STEM teaching–learning sequence on fire ecology, focusing on how preservice and in-service teachers evaluate disciplinary integration, structural coherence, and classroom feasibility. By involving teachers at different stages of professional experience, the study examines patterns in teachers' evaluative judgments and explores differences according to teaching experience and specialization. The findings indicate a high level of perceived disciplinary integration, particularly in the dimensions of argumentation and authenticity, highlighting strengths in the design of the sequence. At the same time, limitations were identified in relation to engineering design and the evaluation of the learning process, pointing to areas for improvement in STEM instructional planning. Statistically significant differences in evaluations were found according to teaching experience, especially in the assessment of the theoretical dimension, with higher ratings from teachers with intermediate experience. Overall, the results illustrate how a structured evaluation rubric can be used to examine the quality of integrated STEM teaching–learning sequences from a teacher perspective, providing empirical evidence on design coherence, disciplinary integration, and classroom applicability.

Keywords: integrated STEM education; teaching–learning sequence; assessment; rubric; fire ecology



Academic Editors: Thomas Roberts,
Stefanie Livers and Holly Plank

Received: 31 December 2025

Revised: 12 February 2026

Accepted: 18 February 2026

Published: 27 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

In recent decades, integrated STEM education—an acronym referring to Science, Technology, Engineering, and Mathematics—has become established as a relevant educational approach for preparing citizens capable of understanding and addressing the scientific and technological challenges of the twenty-first century (Bybee, 2013; Honey et al., 2014). This purpose extends beyond the acquisition of disciplinary knowledge to include the integration of different fields of knowledge in order to foster critical thinking, problem-solving skills, and the application of learning in real-world contexts (Kelley & Knowles, 2016). These aims are reflected in international educational agendas and curricular reforms across different contexts, including Spain, which constitutes the educational setting of the present study (Real Decreto 217/2022, 2022). Nevertheless, the effective implementation of integrated STEM education in classroom practice continues to face significant challenges, including curricular disciplinary fragmentation, limited interdisciplinary preparation of

teachers, a scarcity of appropriate instructional materials, and the lack of clear strategies that support teachers in evaluating and understanding the level of integration achieved in STEM educational proposals (Estévez-Mauriz & Baelo, 2021; Palacios & Aguilera, 2020). Although this study is framed within an integrated STEM perspective, it is also consistent with broader STEAM approaches that emphasize contextualized problem solving, creativity, and socio-environmental relevance, even when the arts are not explicitly foregrounded.

Instructional materials, particularly in the form of teaching–learning sequences, play a key role in the implementation of educational reforms, as they enable teachers to become familiar with new approaches and to develop pedagogical skills and strategies aligned with them. From a teacher education perspective, such materials can function as mediating tools that help teachers articulate and examine key dimensions of integrated STEM design. However, for such materials to be effective, they must be well designed, contextually adapted, and perceived by teachers as relevant and usable in their own classrooms (Thibaut et al., 2018). Accordingly, a recurring challenge in STEM sequences developed either by research groups or by teachers themselves is the absence of evaluation mechanisms that make it possible for teachers to critically examine whether the designed materials achieve an authentic integration of the disciplines involved and to what extent such proposals are perceived as applicable within their professional contexts. Unlike traditional models, in which assessment focuses primarily on the mastery of specific content, the STEM approach requires tools capable of evaluating interdisciplinary coherence, the applicability of knowledge, and alignment with intended learning objectives (Moore, 2014; Roehrig et al., 2021).

To address these challenges, our research group has been working for several years on the design, implementation, and evaluation of integrated instructional materials for the primary education level (Greca et al., 2024; Merino-Fernández et al., 2023; Ortiz-Revilla et al., 2021). More recently, we have also begun to examine the evaluation of the quality of the instructional materials themselves by teachers (Ortiz-Revilla et al., 2024). The results of these studies indicate the need to revise an initial material prototype, which was assessed by users (teachers, students) as partially viable and usable.

Building on this line of work, the present study analyzes a teaching–learning sequence focused on fire ecology, which was previously designed and published as part of a broader instructional proposal for secondary education (Martínez-Martínez & Greca, 2025), with particular attention to how teachers evaluate its design and integration. This topic provides an intuitive context for integrating knowledge from different STEM disciplines to understand fire dynamics in ecosystems, its effects on biodiversity, and its relationship with climate change and environmental management. Accordingly, the aim of this study is to examine, from the perspective of both preservice and in-service teachers, the level of disciplinary integration, structural coherence, and classroom applicability of the STEM teaching–learning sequence. Importantly, this study does not aim to validate the RubeSTEM rubric (the instrument used for doing this assessment, as explained below) nor to assess teachers' professional learning outcomes, but to use the rubric as an analytical lens to examine teachers' evaluations of an integrated STEM teaching–learning sequence. Unlike other studies that focus on the effectiveness of student learning outcomes, this research centers on the evaluation of the design and structuring of the teaching–learning sequence, making it possible to identify strengths and areas for improvement from a teacher perspective, before implementation, in order to better align the didactical material and the eventual teacher training to the real classroom settings. To this end, the following research questions are proposed:

1. To what extent does the STEM teaching–learning sequence on fire ecology integrate the STEM disciplines, according to the evaluations conducted by preservice and in-service teachers?

2. How do preservice and in-service teachers perceive the structural and pedagogical coherence of the teaching–learning sequence?
3. Are there statistically significant differences in the evaluation of the teaching–learning sequence according to teachers’ professional profiles?
4. What adjustments could be made to improve the alignment of the teaching–learning sequence with the pedagogical principles of integrated STEM education?

Through the analysis of these questions, this study seeks to contribute to the transfer of research outcomes into educational practice by providing empirical evidence on teachers’ evaluations of integrated STEM instructional resources.

2. Theoretical Framework

The STEM approach has evolved from a more traditional disciplinary orientation, in which the teaching of science, technology, engineering, and mathematics was addressed separately, toward an integrated perspective that emphasizes the combination of knowledge in applied and authentic contexts (Martín-Páez et al., 2019). This shift responds to the need to educate citizens capable of addressing complex problems that require the integration of multiple disciplines and methodologies. Consequently, contemporary STEM education seeks not only to promote the acquisition of specific knowledge in science, technology, engineering, and mathematics, but also to foster competencies such as problem-solving, critical thinking, and creativity, within a holistic framework of competency-based student development (Ortiz-Revilla et al., 2022).

The implementation of integrated approaches in education in general, and of integrated STEM education in particular, faces multiple barriers at the secondary education level (Asghar et al., 2012; Thibaut et al., 2018). Curricular compartmentalization within educational systems hinders the planning of integrated learning experiences, while insufficient teacher preparation in pedagogical approaches aligned with integration limits teachers’ capacity to design and implement effective instructional strategies (Kurup et al., 2021; Shernoff et al., 2017). Integrated education requires not only a solid command of disciplinary content but also a clear understanding of how different disciplines can be coherently and meaningfully combined. This, in turn, demands ongoing dialogue and collaboration among teachers from different subject areas, who often operate from distinct epistemological perspectives (McComas & Burgin, 2020; Millar, 2020).

In the Spanish context, recent educational legislation has begun to promote educational integration and a more holistic view of learning, including the explicit incorporation of STEM competence from the primary education level onward (Ley Orgánica 3/2020, 2020). Nevertheless, Spanish teachers continue to face a range of challenges when implementing integrated STEM approaches in their classrooms (Arabit García & Prendes Espinosa, 2020; García-Carrillo et al., 2021; Ortiz-Revilla et al., 2023).

The development of instructional proposals has increased in parallel with the significant growth of integrated STEM education (Portillo-Blanco et al., 2024), highlighting the need for appropriate evaluation tools. Due to its nature, traditional assessment approaches that are external to the instructional intervention and focused primarily on the acquisition of discipline-specific knowledge present clear limitations when it comes to assessing the degree of integration and the applicability of knowledge to real-world problem solving (Mejía Pérez, 2012). However, for integrated STEM education to fully achieve its objectives, it is essential that instructional proposals focus not only on knowledge acquisition but also on the interconnection of knowledge and its applicability across diverse contexts.

Accordingly, particular attention must first be paid to the quality of STEM instructional materials themselves, especially with regard to the coherence of disciplinary integration, curricular alignment, classroom applicability, and their potential impact on student learning

(Roehrig et al., 2021). In addition, regardless of the evaluations conducted by researchers, it is critically important to consider teachers’ perceptions of STEM instructional materials, as teachers are ultimately responsible for their implementation in classroom settings.

However, research conducted in this area remains particularly limited. Existing studies, such as those by Domènech-Casal et al. (2019) and Pérez Torres et al. (2021) have focused exclusively on project-based learning approaches within STEM education. Consequently, as noted by Aguilera et al. (2022), there is “a need to move toward greater consensus on rubrics that can serve as guidance for the design of projects and for the evaluation of those already implemented” (p. 17).

Furthermore, as previously indicated, the STEM approach is grounded in the resolution of real-world problems (Roberts et al., 2022). Accordingly, it is both coherent and essential to reflect on the importance of selecting meaningful problem contexts that guide instructional proposals and support the development of significant learning experiences. In this regard, the use of socio-environmental issues in integrated STEM education has been shown to be an effective strategy for fostering integrated learning (Burgess & Buck, 2020). Within this framework, teaching fire ecology constitutes a particularly relevant educational context, as its very nature requires the integration of knowledge from biology, chemistry, physics, mathematics, and geography to analyze the effects of fire on ecosystems, its relationship with climate change, and environmental management strategies (Martínez-Martínez & Greca, 2024; Park et al., 2024). From an educational perspective, the design of integrated instructional proposals on fire ecology offers a valuable opportunity to strengthen STEM education by providing a coherent framework for the integration of scientific content (Martínez-Martínez et al., 2023, 2024). From a STEAM-oriented perspective, socio-environmental problems such as fire ecology also allow for the integration of ethical, communicative, and contextual dimensions that complement scientific and technological reasoning, reinforcing the relevance of interdisciplinary instructional design.

3. Materials and Methods

3.1. Sample

Three groups of teachers participated in the study (Table 1), selected to ensure diversity in terms of professional experience and background in STEM disciplines. The first group consisted of 23 in-service secondary education teachers who attended a professional development course organized by the Teacher Training and Educational Innovation Center of Burgos, entitled Interdisciplinary STEM Learning Situations in Secondary Education. The participants taught different STEM subjects.

Table 1. Sample overview.

Group	Professional Status	Participants	Completed Rubrics
Group 1	In-service teachers	23	23
Group 2	Preservice teachers	28	5
Group 3	Mixed (preservice and in-service)	10	10
Total rubrics analyzed			38

The second group included 28 preservice teachers enrolled in the Master’s degree in Secondary Education Teaching (Compulsory Secondary Education and Upper Secondary Education), Vocational Training, and Language Teaching at the University of Barcelona, specializing in STEM-related disciplines. A distinctive feature of this group was its collaborative approach, as participants worked in teams to complete and submit a total of five

rubrics ($n = 5$). This format encouraged the organization and exchange of ideas, providing a diverse perspective in the analysis of the rubrics. In the case of the preservice teachers (Group 2), the rubric was completed collaboratively in small groups, resulting in five completed rubrics. This approach was intentionally adopted as part of the instructional design of the master's course, where collaborative analysis and consensus-building are key components of professional learning in integrated STEM education. Consequently, the unit of analysis in this group was the jointly constructed evaluative judgment rather than individual responses, which aligns with previous research emphasizing the role of collective reflection in teacher education contexts.

The third group consisted of 10 participants in a summer course on the theory and practice of STEM education organized by the University of the Basque Country. This group included 2 preservice teachers and 8 in-service teachers, all of whom also came from STEM disciplines. This composition made it possible to combine perspectives from initial teacher education and active teaching practice, resulting in a total of 38 responses.

It should be noted that the in-service teachers attended the courses on a voluntary basis, which suggests a certain level of interest in integrated STEM education. All participants were categorized according to their teaching specialization and years of professional experience. Teaching specializations included Biology and Geology ($n = 19$), Physics and Chemistry ($n = 14$), and Technology and Computer Science ($n = 5$). With regard to teaching experience, participants were grouped into those with less than five years of experience ($n = 14$), between five and ten years of experience ($n = 13$), and more than ten years of experience ($n = 11$).

The inclusion of both preservice and in-service teachers was intentional, as the study aims to examine how teachers at different stages of professional development evaluate the design quality of an integrated STEM teaching–learning sequence. Rather than assuming homogeneity, this approach allows for the exploration of convergences and divergences in evaluative judgments across professional trajectories.

3.2. Rubric

Following our previous work (Ortiz-Revilla et al., 2024), we employed an ad hoc adaptation of the RubeSTEM rubric originally developed and validated by Aguilera et al. (2022). This adaptation excluded indicators related to the direct classroom implementation of instructional proposals, with the specific aim of focusing the evaluation on the design and planning phases of the teaching–learning sequence. The adapted rubric comprises a total of 18 indicators, each scored on a four-point ordinal scale ranging from 0 (emerging) to 3 (sophisticated). An analysis of internal consistency yielded a Cronbach's alpha of $\alpha = 0.785$, indicating satisfactory reliability for the purposes of this study.

Rubrics have been widely recognized as effective assessment tools in STEM education, as they allow for the establishment of explicit criteria to analyze the quality of instructional proposals from a structured and transparent perspective. Prior research has shown that rubric-based assessment supports more precise and systematic evaluation processes and can assist teachers in refining instructional designs based on empirical evidence (Reynders et al., 2020). Compared to more general assessment instruments, rubrics offer a differentiated view of multiple design components, thereby facilitating the identification of strengths and areas for improvement in instructional proposals.

Within this framework, RubeSTEM was designed to assess the quality of STEM instructional proposals through two core dimensions: theoretical coherence, represented by the Why dimension, which addresses alignment with the pedagogical principles of integrated STEM education, and practical coherence, represented by the What and How dimension, which focuses on the quality and structural consistency of instructional design

(Aguilera et al., 2022). The original rubric was validated by a panel of 20 experts and tested using 26 STEM proposals developed by 145 preservice primary education teachers in Granada, Spain.

The indicators considered in the present study include: learning objectives (LO), STEM education purposes (STEM-P), the problem addressed by the proposal (PB), disciplinary integration (DI), action deployment (AD), context of implementation and social impact (SI), argumentation (Arg), inquiry (Inq), modeling (Mod), engineering design (ED), impact evaluation (IE), relevance (Rel), authenticity (Aut), evaluation of the process and outcomes (POE), regulation of cooperative work (RCW), opportunities for teacher collaboration (OTC), and opportunities for external agents' involvement (OEA). The additional viability indicator is described separately as a complementary, exploratory dimension.

3.3. Analysis

Following the framework proposed by Aguilera et al. (2022), the indicators were grouped into two dimensions that allow for a comprehensive evaluation of the instructional design. The theoretical dimension, which includes indicators 1 to 3, assesses the extent to which the proposal aligns with the STEM approach, classifying it as pseudo-STEM (scores between 0 and 4) or STEM (scores between 5 and 9). The practical dimension comprises indicators 4 to 17, and its score is calculated as the sum of the individual scores obtained for each of these indicators. This cumulative score reflects the overall quality of the instructional design of the teaching–learning sequence, including aspects such as disciplinary integration, pedagogical coherence, and learning activities. Based on the total score obtained across indicators 4–17, four performance levels were established for this study: emerging (0–13 points), basic (14–27 points), advanced (28–41 points), and sophisticated (42 points).

Finally, in order to determine whether the proposal is truly applicable within each educational context, an additional dimension referred to as viability (V) was incorporated in this study. This dimension included a single indicator focused on the evaluators' perceptions of the feasibility of implementing the proposal in their classrooms, categorized as not viable (0 points), potentially viable (1 point), partially viable (2 points), or viable (3 points).

In addition to generating an overall score that reflects the level of quality and viability of the instructional proposals, the evaluation system made it possible to identify common patterns and potentially problematic aspects in the design of STEM teaching–learning sequences based on teachers' evaluative judgments. The analysis relied exclusively on descriptive and inferential statistical procedures, which provided a systematic basis for interpreting teachers' perceptions across the different dimensions of the rubric. Overall, this process yielded relevant information to inform future refinements of integrated STEM instructional designs and to support evidence-informed discussions in teacher education and professional development contexts.

4. Results

To address the research objectives and questions posed, two subsections of results are presented below, corresponding to the descriptive and inferential statistical analyses conducted on the data obtained from the participating teachers.

4.1. Descriptive Statistics of the Evaluated Indicators

To analyze the responses obtained, descriptive statistics were first calculated for each indicator and dimension of the applied rubric. Table 2 presents the descriptive statistics for each indicator, including the mean, standard deviation, minimum and maximum values,

and the total number of valid responses, together with the frequency (f) distribution of scores across the four performance levels

Table 2. Descriptive statistics by indicator.

Dimension	Indicator	N	0 (f)	1(f)	2 (f)	3 (f)	Minimum	Maximum	Mean	Standard Deviation
Theoretical	LO	38	0	0	19	19	2	3	2.50	0.507
	STEM-P	38	2	5	24	7	0	3	1.95	0.733
	PB	38	0	5	20	13	1	3	2.21	0.664
Practical	DI	38	2	2	15	19	0	3	2.34	0.815
	AD	38	0	4	27	7	1	3	2.08	0.539
	SI	38	9	2	15	12	0	3	1.79	1.143
	Arg	38	0	0	8	30	2	3	2.79	0.413
	Inq	38	1	4	19	14	0	3	2.21	0.741
	Mod	38	1	7	8	22	0	3	2.34	0.878
	ED	38	9	11	11	7	0	3	1.42	1.056
	IE	38	4	0	15	19	0	3	2.29	0.927
	Rel	38	0	2	17	19	1	3	2.45	0.602
	Aut	38	0	4	10	24	1	3	2.53	0.687
	POE	38	10	11	11	6	0	3	1.34	1.047
	RCW	38	4	11	12	11	0	3	1.79	0.991
	OTC	38	0	6	13	19	1	3	2.34	0.745
	OEA	38	0	4	21	13	1	3	2.24	0.634
Viability	V	38	1	7	16	14	0	3	2.13	0.811

These results reveal that the items with the highest mean scores were Argumentation ($M = 2.79$, $SD = 0.413$) and Authenticity ($M = 2.53$, $SD = 0.687$), with a high concentration of ratings at the upper levels of the scale, whereas the lowest-scoring items were Process evaluation ($M = 1.34$, $SD = 1.047$) and Engineering design ($M = 1.42$, $SD = 1.056$), which also showed greater dispersion of scores across the lower and intermediate levels. This pattern indicates that certain aspects of the teaching–learning sequence exhibited more homogeneous and higher performance, while others showed greater variability and lower scores on the scale used.

Subsequently, descriptive statistics were calculated for the three overall dimensions of the rubric: the theoretical dimension (D_Theo), the practical dimension (D_Prac), and the viability dimension (D_Via). These results are presented in Table 3:

Table 3. Descriptive statistics by dimension.

Dimension	N	Minimum	Maximum	Mean	Standard Deviation
D_Theo	38	4	9	6.66	1.381
D_Prac	38	18	39	29.95	5.647
D_Via	38	0	3	2.13	0.811

Descriptive statistics were examined separately for each dimension, taking into account their distinct score ranges and measurement scales. In the theoretical dimension

(range 0–9), the mean score ($M = 6.66$, $SD = 1.381$) places the teaching–learning sequence within the range corresponding to a STEM proposal, according to the scoring criteria established by Aguilera et al. (2022), this value falls within the range corresponding to a STEM proposal. In the practical dimension (range 0–42), the mean score obtained ($M = 29.95$, $SD = 5.647$) corresponds to an advanced level of instructional design quality, reflecting a generally positive evaluation of the structure and integration of the sequence.

With respect to the viability dimension (ordinal scale 0–3), the mean score ($M = 2.13$, $SD = 0.811$) indicates that the sequence was perceived, on average, as lying between partially viable and fully viable. Given the ordinal nature of this dimension, these values should be interpreted as descriptive tendencies rather than interval-based measurements. However, the dispersion of scores in this dimension suggests greater heterogeneity in teachers' perceptions regarding feasibility, pointing to differences in how easily the sequence is perceived as implementable across classroom contexts.

From a descriptive perspective, differences were also observed across groups defined by teaching experience. In the practical dimension, teachers with less than five years of experience obtained the highest mean score ($M = 31.21$, $SD = 7.266$), closely followed by those with five to ten years of experience ($M = 31.08$, $SD = 3.013$), and teachers with more than ten years of experience ($M = 27.00$, $SD = 5.00$). Regarding the viability dimension, the highest mean score was observed in the group with five to ten years of experience ($M = 2.31$, $SD = 0.751$), followed by the group with less than five years ($M = 2.29$, $SD = 0.611$) and the group with more than ten years of experience ($M = 1.73$, $SD = 1.009$).

From a descriptive perspective, variations in mean scores were observed across teaching specializations. In the theoretical dimension, the highest mean score was obtained by teachers specializing in Technology and Computer Science ($M = 7.00$, $SD = 2.00$), followed by Biology and Geology ($M = 6.74$, $SD = 1.447$) and Physics and Chemistry ($M = 6.43$, $SD = 1.089$). A similar pattern was observed in the practical dimension, where Technology and Computer Science teachers again showed the highest mean scores ($M = 30.40$, $SD = 7.503$), followed by Biology and Geology ($M = 30.32$, $SD = 5.879$) and Physics and Chemistry ($M = 29.29$, $SD = 4.983$). In contrast, in the viability dimension, the highest mean score was observed among Biology and Geology teachers ($M = 2.26$, $SD = 0.733$), followed by Physics and Chemistry ($M = 2.07$, $SD = 0.829$) and Technology and Computer Science ($M = 1.80$, $SD = 1.095$).

4.2. Inferential Statistics of the Evaluated Indicators

To assess the normality of the data distribution, the Kolmogorov–Smirnov test was applied. The results (Table 4) indicate that none of the dimensions followed a normal distribution, as the obtained p -values were consistently significant ($p < 0.05$). This justified the use of nonparametric tests for subsequent analyses. Specifically, the Kruskal–Wallis H test was employed to compare groups according to years of teaching experience and teaching specialization.

Table 4. Kolmogorov–Smirnov test by dimension.

	D_Theo	D_Prac	D_V
N	38	38	38
Test statistic	0.203	0.153	0.226
Asymp. Sig. (two-tailed) ¹	<0.001	0.025	<0.001

¹ Lilliefors significance correction.

When controlling for years of teaching experience, the results of the Kruskal–Wallis H test (Table 5) did not reveal statistically significant differences in the practical dimension

($H = 5.268, p = 0.072$). Although the associated effect size was small to moderate ($\epsilon^2 = 0.081$), this suggests that any observed differences between experience groups were modest and did not reach sufficient magnitude to support robust practical interpretations.

Table 5. Kruskal–Wallis test by years of teaching experience.

	D_Theo	D_Prac	D_V
Kruskal–Wallis H	10.339	5.268	2.943
df	2	2	2
Asymp. Sig.	0.006	0.072	0.023

Statistically significant differences were observed in the theoretical dimension ($H = 10.339, p = 0.006$), with a moderate effect size ($\epsilon^2 = 0.159$), indicating a meaningful association between years of teaching experience and teachers' theoretical evaluations of the teaching–learning sequence. In the viability dimension, statistically significant differences were also found ($H = 2.943, p = 0.023$); however, the effect size was small ($\epsilon^2 = 0.045$), indicating that although differences reached statistical significance, their practical relevance was limited. These results should therefore be interpreted with caution, as small effect sizes suggest only minor variations that may be influenced by contextual or sampling-related factors rather than substantive differences between experience groups.

To complement the inferential analysis and provide a distributional perspective consistent with the use of nonparametric tests, box-and-whisker plots were employed. This graphical representation displays medians, interquartile ranges, and overall score dispersion without assuming normality, making it appropriate for non-normally distributed data. Figure 1 illustrates the distribution of scores across experience groups in the practical and viability dimensions.

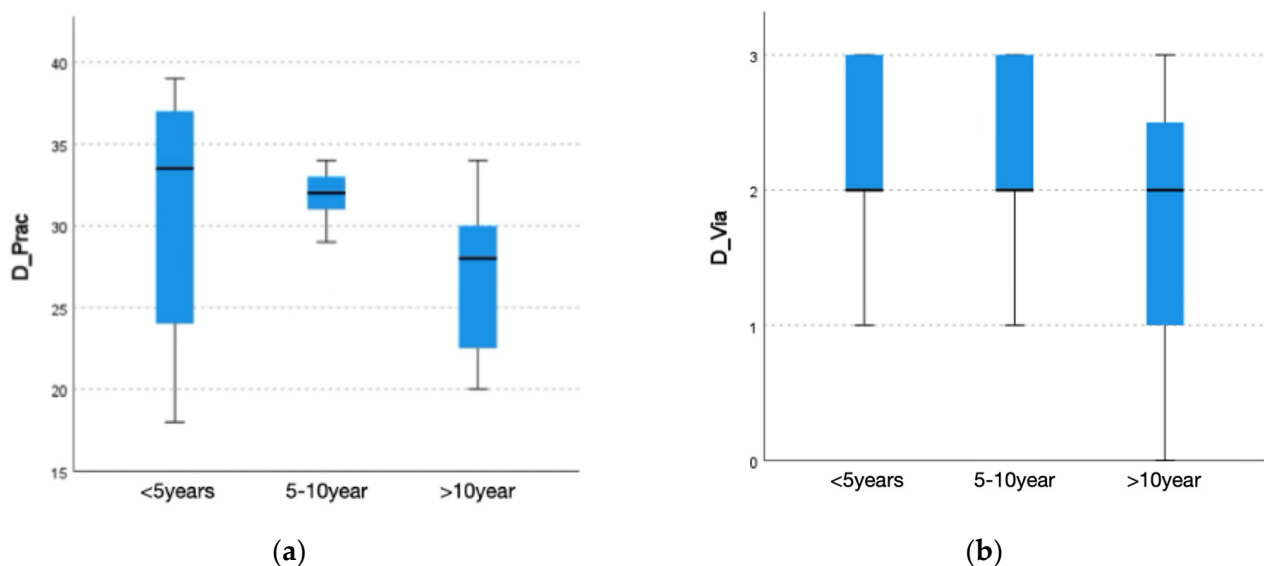


Figure 1. Box-and-whisker plots showing the distribution of scores for the practical dimension (a) and the viability dimension (b) according to years of teaching experience (<5 years, 5–10 years, >10 years). The plots display median values, interquartile ranges, and overall score dispersion across experience groups.

Post hoc pairwise comparisons using Dunn's test with Bonferroni correction indicated that the theoretical dimension showed statistically significant differences (adjusted $p = 0.005$) between teachers with more than ten years of experience ($M = 5.64, SD = 1.027$)

and those with five to ten years of experience ($M = 7.31$, $SD = 1.251$), in favor of the latter. By contrast, no statistically significant differences were found between teachers with more than ten years of experience and those with less than five years of experience (adjusted $p = 0.066$), although the latter group obtained higher scores. Similarly, no significant differences were observed between teachers with less than five years of experience and those with five to ten years of experience (adjusted $p = 1.000$), again with higher scores obtained by the latter group. Figure 2 presents the node plot and the corresponding box-and-whisker plots for these comparisons.

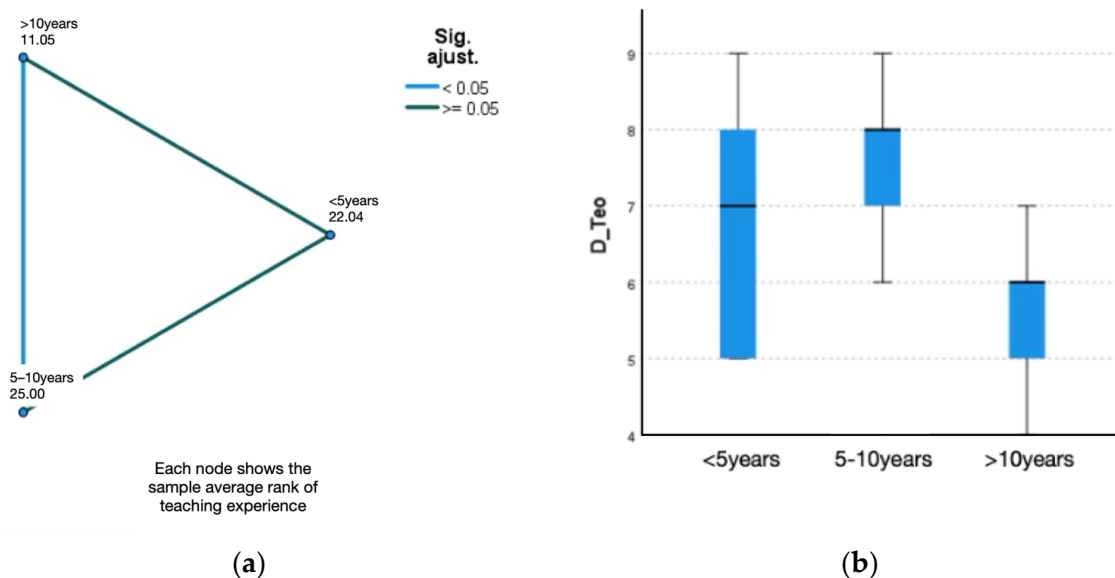


Figure 2. Node plot (a) and box-and-whisker plot (b) illustrating the distribution of scores for the theoretical dimension by years of teaching experience.

When controlling for teaching specialization, the Kruskal–Wallis H test did not detect statistically significant differences in any of the evaluated dimensions (Table 6), indicating that perceptions of the teaching–learning sequence did not vary substantially across the different teaching specializations considered. In all cases, the associated effect sizes were negligible ($\epsilon^2 \leq 0.021$), indicating minimal practical relevance of the observed differences.

Table 6. Kruskal–Wallis test by teacher specialization.

	D_Theo	D_Prac	D_V
Kruskal–Wallis H	1.348	0.721	0.958
df	2	2	2
Asymp. Sig.	0.510	0.697	0.619

When controlling for teaching specialization, the Kruskal–Wallis H test did not detect statistically significant differences in any of the evaluated dimensions (Table 6). This indicates that the distribution of scores across specialization groups did not differ significantly in theoretical, practical, or viability evaluations. To provide a distributional perspective consistent with the nonparametric analysis, Figure 3 presents box-and-whisker plots illustrating the median values and dispersion patterns across specialization groups in the three dimensions.

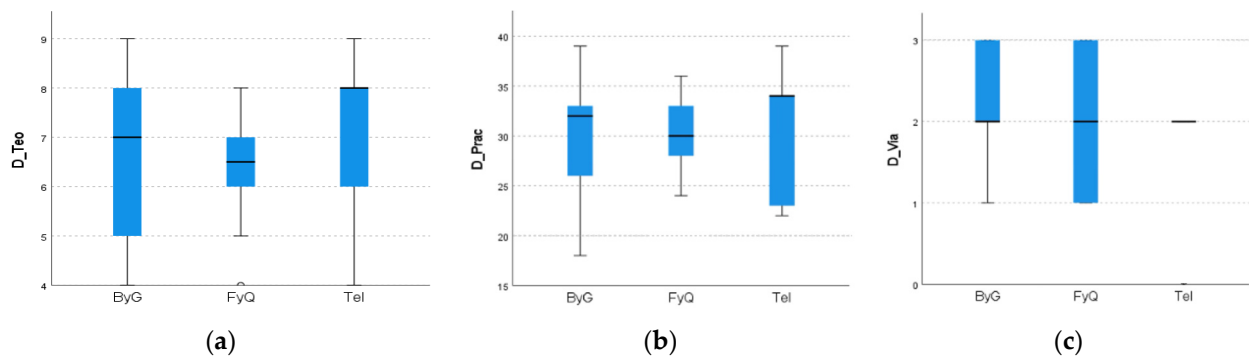


Figure 3. Box-and-whisker plots showing the distribution of scores by teaching specialization for (a) the theoretical dimension, (b) the practical dimension and (c) the viability dimension. Teaching specializations include Biology and Geology (ByG), Physics and Chemistry (FyQ), and Technology and Computer Science (Tel).

5. Discussion

The results obtained allow for reflection on the quality and applicability of the STEM teaching–learning sequence, as examined through teachers’ evaluations using a structured analytic rubric, as well as on the challenges and opportunities associated with its implementation. The mean scores in the theoretical and practical dimensions indicate that teachers perceive the fire ecology STEM sequence as an advanced STEM proposal; that is, it is highly valued in terms of both theoretical and practical quality. These results improve upon those obtained previously in an initial prototype of instructional material for primary education, in which teachers rated the proposal as basic pseudo-STEM (Ortiz-Revilla et al., 2024). Beyond informing the quality of the teaching–learning sequence, these results also shed light on how teachers at different stages of their professional trajectories evaluate key dimensions of integrated STEM design.

From the teachers’ perspective, the teaching–learning sequence was particularly positively valued for its emphasis on critical thinking and problem-solving, aspects reflected in the highest-rated items of Argumentation and Authenticity. These competencies constitute two fundamental pillars of STEM education (English, 2023; Tan et al., 2023). Nevertheless, opportunities for improvement were identified in specific aspects such as the integration of engineering design and the strengthening of strategies for assessing the learning process, which correspond to the lowest-rated items. In particular, the low score obtained for the engineering design item suggests that this dimension may require the reformulation of some activities to achieve greater alignment in this respect. However, it is also possible that these results are related to the limited knowledge that both preservice and in-service teachers often have regarding the nature of engineering and engineering design methodology, an issue widely reported in the literature (Carbajo-Barbero et al., 2021; Chai et al., 2020; Deniz et al., 2020). This interpretation is further supported by researchers’ observations recorded in the field notes during work with the different groups of teachers. The low score obtained for the evaluation of the learning process reinforces a need previously identified in the literature: to make assessment strategies explicit through tools and systems capable of capturing the complexity of integrated STEM teaching–learning processes (Saxton et al., 2014), with rubrics being among the most frequently used instruments in this approach (Huang & Jong, 2020; Karakaya & Yilmaz, 2022). From a teacher education perspective, these findings highlight the importance of making engineering design processes and assessment strategies more explicit within integrated STEM instructional designs and the criteria used to evaluate them.

Overall, these findings suggest that, although the sequence is positively valued—particularly in its theoretical and practical quality—there may be structural or methodological difficulties that could affect its implementation.

A key limitation of this study relates to participant self-selection. In-service teachers took part in the study on a voluntary basis through professional development activities focused on integrated STEM education, which suggests a pre-existing interest in innovative or interdisciplinary approaches. This self-selection may have contributed to the generally positive evaluations observed and limits the generalizability of the findings to the broader population of secondary school teachers, particularly those with lower levels of engagement with STEM integration. In addition, a further methodological limitation concerns the collaborative completion of the rubric by preservice teachers in Group 2. While this approach was intentionally adopted to promote discussion and shared reflection, it may have reduced the independence of individual judgments. Consequently, results involving this group should be interpreted with caution, particularly in inferential comparisons with individually completed evaluations. In addition, the sample did not include teachers specializing in Mathematics, which restricts the representativeness of the disciplinary perspective within the STEM framework. The absence of this specialization limits the extent to which the findings can be generalized across all STEM subject areas and may have influenced the patterns observed in the evaluation of disciplinary integration and classroom viability. Future research with larger, more balanced, and more diverse samples—including all STEM specializations—is needed to examine the robustness of the findings and to determine whether the small effect sizes observed in some dimensions reflect meaningful differences in practice or are primarily attributable to contextual and sample-related factors.

Regarding differences related to teaching experience, the results indicate that years of teaching experience have a significant impact on teachers' theoretical perceptions of the sequence. Significant differences were found in the theoretical dimension between teachers with more than ten years of experience and those with five to ten years of experience. However, no statistically significant differences were found between teachers with more than ten years of experience and those with fewer than five years, nor between the latter and the intermediate group.

Teachers with more than ten years of experience tended to evaluate disciplinary integration and structural coherence more critically. This pattern is consistent with previous studies in which less experienced teachers were found to have a less robust understanding of integration and therefore to assess it less critically (Ortiz-Revilla et al., 2023), a trend that also emerged in the present study among teachers with fewer years of experience. These results suggest that professional experience influences the identification of methodological and structural aspects, as well as evaluative perspectives that may vary according to professional experience (Tao, 2019). Rather than being interpreted solely as differences in individual attitudes, these patterns may reflect distinct stages of professional experience, in which teachers progressively apply more differentiated criteria to judge disciplinary integration and pedagogical coherence in STEM designs.

However, it is also important to consider the existence of negative attitudes toward new integrated educational approaches among some highly experienced teachers approaching the end of their careers, as reported in previous research (Thibaut et al., 2019). Although no statistically significant differences were found, a similar trend was observed in the practical dimension, where teachers with low and intermediate levels of experience assigned very similar scores, both higher than those of the most experienced group. On the other hand, no statistically significant differences were found in teachers' perceptions of the teaching–learning sequence according to teaching specialization, suggesting that the educational proposal is perceived as suitable across different STEM disciplines. The proximity of the

scores obtained by the three groups of specialists in the theoretical and practical dimensions indicates a relative homogeneity in their evaluations. This apparent convergence makes it difficult to identify consistent differential patterns, which is why we consider it premature to draw definitive conclusions in this regard.

Overall, the results show that, according to teachers' perceptions, the teaching–learning sequence demonstrates an advanced level of disciplinary integration, indicating that it was perceived as being designed with a structure that enables meaningful connections across STEM disciplines. This represents a response to one of the most recurrent demands in the literature on integrated STEM education (Roehrig et al., 2021). In this respect, the thematic focus on fire ecology appears to have played a key role, as it constitutes a topic that allows for intuitive integration of knowledge from different STEM disciplines. Beyond a STEM framework, this type of socio-environmental context also aligns with STEAM perspectives by opening spaces for discussion of ethical decision-making, risk communication, and societal impacts, even if these dimensions were not explicitly evaluated in the present study. Caution is therefore required when extrapolating these results to STEM contexts with substantially different epistemic or pedagogical foundations.

Despite these generally positive results, variability was observed in the viability dimension, indicating that not all teachers perceive the implementation of the sequence as equally feasible. In this regard, no statistically significant differences were found in the viability dimension according to years of teaching experience, nor was a clear correlation identified between teaching experience and viability scores. The highest viability scores were obtained by teachers with intermediate experience (five to ten years). For the most experienced teachers, the lower viability ratings may be interpreted as being associated with a more conservative stance. Previous studies have suggested that veteran teachers, despite their extensive pedagogical background, tend to exhibit greater resistance to change, particularly when it involves substantial transformations of established practices, as is the case with integrated approaches such as STEM education (Morgado et al., 2021; Weinberg et al., 2021). Conversely, novice teachers, although often more open to innovation, may experience greater difficulty in perceiving complex proposals as viable due to limited training in the design of integrated approaches and a lack of teaching experience (Al Salami et al., 2017). Taken together, these findings suggest that perceptions of viability are not fixed but contingent on teachers' experience and perception, prior training, and access to sustained professional support when engaging with integrated STEM proposals. From an applied perspective, perceptions of limited viability highlight the importance of providing concrete forms of support for teachers, such as structured planning templates, adaptable timelines, access to shared instructional resources, and opportunities for collaborative planning with colleagues. These forms of support may reduce perceived implementation barriers and facilitate the appropriation of integrated STEM proposals in diverse classroom contexts.

No statistically significant differences were found in the viability dimension according to teaching specialization either. In this case, the highest scores were assigned by Biology and Geology teachers, followed closely by Physics and Chemistry teachers, with lower scores assigned by Technology and Computer Science teachers. This pattern may be interpreted considering the thematic focus of the teaching–learning sequence analyzed in this study. It is reasonable to assume that Biology, Geology, Physics and Chemistry teachers are more familiar with a sequence based on fire ecology, given its strong emphasis on scientific content directly related to their disciplines. It is possible that the sequence did not sufficiently foreground elements more closely aligned with Technology and Computer Science, despite the inclusion of activities related to information and communication technologies, such as the use of drones, simulation software, and satellite resource analysis. Nevertheless, previous studies have indicated that teachers with technical backgrounds

may show greater resistance to the implementation of innovative methodologies (Rigler, 2016). Taken together, these findings highlight the challenges inherent in implementing classroom innovations, particularly in the case of multi-, inter-, and transdisciplinary proposals (Gresnigt et al., 2014). This is consistent with prior research emphasizing the need to adapt interdisciplinary proposals to real classroom conditions and to the structural constraints of the educational systems in which they are implemented (Holmlund et al., 2016; Murray et al., 2020). From this perspective, evaluation tools such as the one used here can function not only as assessment instruments but also as analytic frameworks that help preservice and in-service teachers articulate, examine, and compare key dimensions of integrated STEM instructional designs.

6. Conclusions

The results obtained in this study allow for drawing conclusions regarding the evaluation of a STEM teaching–learning sequence from teachers’ perspectives, with relevant and novel educational implications for research on integrated STEM education. In addition, although statistically significant differences were identified in some dimensions according to teaching experience, the associated effect sizes indicate that these differences were moderate or small. This suggests that, in practical terms, variations in teachers’ evaluations should be interpreted with caution, as they may reflect nuanced tendencies rather than substantial or systematic differences between professional groups. Future research with larger and more balanced samples is therefore needed to further examine the practical significance of these effects and to explore how contextual and professional factors shape teachers’ evaluative judgments. Against this backdrop, the findings highlight that the availability of instructional materials perceived as high quality is associated with more favorable teacher evaluations and perceptions of classroom applicability. In this sense, high-quality instructional materials contribute to advancing processes of educational renewal and innovation.

Overall, the analysis presented in this paper made it possible to identify strengths and areas for improvement in the design of the STEM teaching–learning sequence, providing an analytical framework to assess its disciplinary integration, structural coherence, and classroom viability.

Based on these findings, a set of improvements is proposed to optimize the STEM teaching–learning sequence and enhance its alignment with quality principles in integrated STEM education, which may also be extrapolated to the design of future instructional proposals. First, greater emphasis should be placed on strengthening the integration of engineering design within the instructional sequence by promoting activities that enable students to apply engineering principles to the resolution of contextualized STEM problems (Roberts et al., 2022; Rodrigues-Silva et al., 2023), as well as specific training and/or materials for teachers about this topic. Second, particular attention should be given to the assessment of the learning process through strategies that address not only the evaluation of final products, but also the monitoring of competency development throughout the unit and its activities (Grangeat et al., 2021). The use of detailed rubrics, along with the incorporation of self-assessment and peer-assessment mechanisms, may contribute to this aim.

In addition, efforts should be directed toward greater flexibility in instructional planning to allow adaptations to different educational contexts, thereby enhancing the viability of the teaching–learning sequence across diverse teaching environments. This is especially feasible given the versatility of the thematic focus around which the sequence is structured. From a broader educational perspective, teacher education in integrated education strategies represents an important contextual factor in ensuring more effective implementation

of this type of instructional proposal in classroom practice (Lantau et al., 2020). Moreover, to ensure effective implementation, such training should be initiated and sustained within official teacher education programs (Castro-Rodríguez & Montoro, 2021).

In conclusion, this study highlights the value of using an established analytic rubric such as RubeSTEM to examine teachers' evaluations of the design quality of an integrated STEM teaching–learning sequence. The analysis made it possible to identify specific strengths of the sequence—particularly in terms of disciplinary integration, argumentation, and authenticity—as well as areas requiring further refinement, such as engineering design and the assessment of learning processes. Rather than validating the instrument itself, the findings illustrate how structured evaluative frameworks can support systematic reflection on STEM instructional design from a teacher perspective. In this sense, the results provide empirically grounded insights to inform the iterative improvement of integrated STEM teaching–learning sequences and to guide discussions in teacher education contexts about the characteristics of high-quality STEM instructional designs.

Author Contributions: Conceptualization, I.M.G.; methodology, V.M.-M. and J.O.-R.; software, V.M.-M. and J.O.-R.; formal analysis, V.M.-M. and J.O.-R.; investigation, V.M.-M.; resources, V.M.-M.; data curation, V.M.-M.; writing—original draft preparation, V.M.-M.; writing—review and editing, J.O.-R. and I.M.G.; visualization, V.M.-M.; supervision, I.M.G.; project administration, I.M.G.; funding acquisition, I.M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Innovation of the Government of Spain. (Grant number PID2020-118010RB-I00).

Institutional Review Board Statement: Ethical approval for the study was obtained from the Ethics Committee of the University of Burgos (Approval code IR 33/2023, with grant on 24 October 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Aguilera, D., García-Yeguas, A., Palacios, F. J. P., & Vílchez-González, J. M. (2022). Diseño y validación de una rúbrica para la evaluación de propuestas didácticas STEM (RUBESTEM). *Revista Interuniversitaria de Formación del Profesorado. Continuación de la antigua Revista de Escuelas Normales*, 97(36.1), 11–34. [\[CrossRef\]](#)
- Al Salami, M. K., Makela, C. J., & De Miranda, M. A. (2017). Assessing changes in teachers' attitudes toward interdisciplinary STEM teaching. *International Journal of Technology and Design Education*, 27(1), 63–88. [\[CrossRef\]](#)
- Arabit García, J., & Prendes Espinosa, M. P. (2020). Metodologías y tecnologías para enseñar STEM en educación primaria: Análisis de necesidades. *Pixel-Bit, Revista de Medios y Educación*, 57, 107–128. [\[CrossRef\]](#)
- Asghar, A., Ellington, R., Rice, E., Johnson, F., & Prime, G. M. (2012). Supporting STEM education in secondary science contexts. *Interdisciplinary Journal of Problem-Based Learning*, 6(2), 4. [\[CrossRef\]](#)
- Burgess, A., & Buck, G. A. (2020). Inquiring into environmental STEM: Striving for an engaging inquiry-based E-STEM experience for pre-service teachers. In V. L. Akerson, & G. A. Buck (Eds.), *Critical questions in STEM education* (Vol. 51, pp. 61–84). Springer International Publishing. [\[CrossRef\]](#)
- Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities*. NSTA Press.
- Carbajo-Barbero, R. M., Ortiz-Revilla, J., & Greca, I. M. (2021). Comprensión de algunas características de la ingeniería en maestros en formación. *Revista de Enseñanza de la Física*, 33, 135–144.
- Castro-Rodríguez, E., & Montoro, A. B. (2021). Educación STEM y formación del profesorado de Primaria en España. *Revista de Educación*, 393, 353–378. [\[CrossRef\]](#)
- Chai, C. S., Rahmawati, Y., & Jong, M. S.-Y. (2020). Indonesian science, mathematics, and engineering preservice teachers' experiences in STEM-TPACK design-based learning. *Sustainability*, 12(21), 9050. [\[CrossRef\]](#)
- Deniz, H., Kaya, E., Yesilyurt, E., & Trabia, M. (2020). The influence of an engineering design experience on elementary teachers' nature of engineering views. *International Journal of Technology and Design Education*, 30(4), 635–656. [\[CrossRef\]](#)

- Domènech-Casal, J., Lope, S., & Mora, L. (2019). Qué proyectos STEM diseña y qué dificultades expresa el profesorado de secundaria sobre Aprendizaje Basado en Proyectos. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 16(2), 2203. [CrossRef]
- English, L. D. (2023). Ways of thinking in STEM-based problem solving. *ZDM—Mathematics Education*, 55(7), 1219–1230. [CrossRef]
- Estévez-Mauriz, L., & Baelo, R. (2021). How to evaluate the STEM curriculum in Spain? *Mathematics*, 9(3), 236. [CrossRef]
- García-Carrillo, C., Greca, I. M., & Fernández-Hawrylak, M. (2021). Teacher perspectives on teaching the STEM approach to educational coding and robotics in primary education. *Education Sciences*, 11(2), 64. [CrossRef]
- Grangeat, M., Harrison, C., & Dolin, J. (2021). Exploring assessment in STEM inquiry learning classrooms. *International Journal of Science Education*, 43(3), 345–361. [CrossRef]
- Greca, I. M., Ortiz-Revilla, J., Radu Bogan, T., & García-Terceño, E. M. (2024). Mucho más que unidades: Propuestas didácticas integradas para educación primaria. In *31 encuentros internacionales de didáctica de las ciencias experimentales: Hacia una educación científica alineada con la Agenda 2030*. Servicio de Publicaciones e Imagen Institucional de la Universidad de Burgos.
- Gresnigt, R., Taconis, R., Van Keulen, H., Gravemeijer, K., & Baartman, L. (2014). Promoting science and technology in primary education: A review of integrated curricula. *Studies in Science Education*, 50(1), 47–84. [CrossRef]
- Holmlund, T., Lesseig, K., Slavitt, D., & Seidel, R. (2016). Supporting middle school teachers' implementation of STEM design challenges. *School Science and Mathematics*, 116, 177–188. [CrossRef]
- Honey, M. A., Pearson, G., & Schweingruber, H. (2014). *STEM integration in K-12 education: Status, prospects, and an agenda for research* (p. 165). The National Academies Press. [CrossRef]
- Huang, B., & Jong, M. S. (2020). Developing a generic rubric for evaluating students' work in STEM education. In *2020 International Symposium on Educational Technology (ISET)* (pp. 210–213). IEEE. [CrossRef]
- Karakaya, F., & Yilmaz, M. (2022). Teachers views on assessment and evaluation methods in STEM education: A science course example. *Journal of Pedagogical Research*, 6(2), 61–71. [CrossRef]
- Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education*, 3(1), 11. [CrossRef]
- Kurup, P. M., Yang, Y., Li, X., & Dong, Y. (2021). Interdisciplinary and integrated STEM. *Encyclopedia*, 1(4), 1192–1199. [CrossRef]
- Lantau, J.-M., Bracke, M., Bock, W., & Capraro, P. (2020). The design of a successful teacher training to promote interdisciplinary STEM modelling projects. In G. A. Stillman, G. Kaiser, & C. E. Lampen (Eds.), *Mathematical modelling education and sense-making* (pp. 455–465). Springer International Publishing. [CrossRef]
- Ley Orgánica 3/2020. (2020). *Ley orgánica 3/2020, de 29 de diciembre, por la que se modifica la ley orgánica 2/2006, de 3 de mayo, de educación*. Pub. L. No. Ley Orgánica 3/2020, BOE-A-2020-17264 122868. Available online: <https://www.boe.es/eli/es/lo/2020/12/29/3> (accessed on 11 February 2026).
- Martínez-Martínez, V., & Greca, I. M. (2024). Theoretical framework for a didactic nature-of-STEM sequence. In D. Ortega-Sánchez (Ed.), *Transdisciplinarity in citizenship education: Challenges, advances, and research proposals* (pp. 23–46). Springer Nature. [CrossRef]
- Martínez-Martínez, V., & Greca, I. M. (2025). *Incendios (cada vez más) impredecibles. Una propuesta interdisciplinar desde la Naturaleza de STEM*. Servicio de Publicaciones e Imagen Institucional de la Universidad de Burgos. [CrossRef]
- Martínez-Martínez, V., Ortiz-Revilla, J., Greca, I., Merlin, A., Sammaritano, M., López, M., & Molina, R. (2023, August 9–11). *Ecología del Fuego: Una herramienta para comprobar los conocimientos acerca de Naturaleza de STEM*. HPS&ST em Tempos de Negação Científica. 5ª Conferência Latinoamericana do International History, Philosophy and Science Teaching Group, Porto Alegre, Brazil. Available online: https://www.ufrgs.br/ihpstla2023/wp-content/uploads/2023/10/Anais_IHPST_2023.pdf (accessed on 11 February 2026).
- Martínez-Martínez, V., Ortiz-Revilla, J., & Greca, I. M. G. (2024). Estudio previo en el diseño y validación de un instrumento para identificar el conocimiento sobre naturaleza de STEM. *Ápice. Revista de Educación Científica*, 8(1), 85–100. [CrossRef]
- Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vilchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education*, 103(4), 799–822. [CrossRef]
- McComas, W. F., & Burgin, S. R. (2020). A critique of “STEM” education: Revolution-in-the-making, passing fad, or instructional imperative? *Science & Education*, 29(4), 805–829. [CrossRef]
- Mejía Pérez, O. (2012). De la evaluación tradicional a una nueva evaluación basada en competencias. *Revista Electrónica Educare*, 16(1), 27–46. [CrossRef]
- Merino-Fernández, M. Á., Ortiz-Revilla, J., & Greca, I. M. (2023). Diseño universal para el aprendizaje en educación STEAM* integrada: Una experiencia en educación primaria. *Revista de Enseñanza de la Física*, 35(2), 223–235. [CrossRef]
- Millar, V. (2020). Trends, issues and possibilities for an interdisciplinary STEM curriculum. *Science & Education*, 29(4), 929–948. [CrossRef]
- Moore, A. (2014). *Understanding the School Curriculum: Theory, politics and principles*. Routledge. [CrossRef]
- Morgado, J. C., Lencastre, J. A., Freires, T., & Bento, M. (2021). Smart education as empowerment: Outlining veteran teachers' training to promote digital migration. *Technology, Knowledge and Learning*, 26(4), 897–916. [CrossRef]

- Murray, J., Cuen Paxson, L., Seo, S., & Beattie, M. (2020, June 22–26). *STEM-oriented alliance for research (SOAR): An educational model for interdisciplinary project-based learning*. 2020 ASEE Virtual Annual Conference Content Access, Virtual Online. [CrossRef]
- Ortiz-Revilla, J., Greca, I. M., & Arriassecq, I. (2022). A theoretical framework for integrated STEM education. *Science & Education*, 31(2), 383–404. [CrossRef]
- Ortiz-Revilla, J., Greca, I. M., & Meneses-Villagr , J.- . (2021). Effects of an integrated STEAM approach on the development of competence in primary education students (Efectos de una propuesta STEAM integrada en el desarrollo competencial del alumnado de Educaci n Primaria). *Journal for the Study of Education and Development*, 44(4), 838–870. [CrossRef]
- Ortiz-Revilla, J., Ruiz-Mart n,  ., & Greca, I. M. (2023). Conceptions and attitudes of pre-school and primary school teachers towards STEAM education in Spain. *Education Sciences*, 13(4), 377. [CrossRef]
- Ortiz-Revilla, J., Sanz-Camarero, R., & Greca, I. M. (2024). Investigaci n de dise o y educaci n STEAM integrada: Experiencias pioneras y reflexiones. In *31 Encuentros internacionales de did ctica de las ciencias experimentales: Hacia una educaci n cient fica alineada con la agenda 2030*. Servicio de Publicaciones e Imagen Institucional.
- Palacios, F. J. P., & Aguilera, D. (2020). Ciencia-Tecnolog a-Sociedad vs. STEM:  evoluci n, revoluci n o disyunci n?  pice. *Revista de Educaci n Cient fica*, 4(1), 1–15. [CrossRef]
- Park, W., Lim, I., & Song, J. (2024). Exploring the intersection of disasters and science education with preservice science teachers through a disaster case study. *Cultural Studies of Science Education*, 19(4), 593–622. [CrossRef]
- P rez Torres, M., Couso, D., & M rquez, C. (2021).  C mo dise ar un buen proyecto STEM? Identificaci n de tensiones en la co-construcci n de una r brica para su mejora. *Revista Eureka sobre Ense anza y Divulgaci n de las Ciencias*, 18(1), 1301. [CrossRef]
- Portillo-Blanco, A., Deprez, H., De Cock, M., Guisasola, J., & Zuza, K. (2024). A systematic literature review of integrated STEM education: Uncovering consensus and diversity in principles and characteristics. *Education Sciences*, 14(9), 1028. [CrossRef]
- Real Decreto 217/2022. (2022). *Real decreto 217/2022, de 29 de marzo, por el que se establece la ordenaci n y las ense anzas m nimas de la educaci n secundaria obligatoria*. Pub. L. No. Real Decreto 217/2022, BOE-A-2022-4975 41571. Available online: <https://www.boe.es/eli/es/rd/2022/03/29/217> (accessed on 11 February 2026).
- Reynders, G., Lantz, J., Ruder, S. M., Stanford, C. L., & Cole, R. S. (2020). Rubrics to assess critical thinking and information processing in undergraduate STEM courses. *International Journal of STEM Education*, 7(1), 9. [CrossRef]
- Rigler, K. L. (2016). Educators’ resistance to the technology and engineering education transition. *Journal of Technology Studies*, 42(1), 42–53. [CrossRef]
- Roberts, T., Maiorca, C., Jackson, C., & Mohr-Schroeder, M. (2022). Integrated STEM as problem-solving practices. *Investigations in Mathematics Learning*, 14(1), 1–13. [CrossRef]
- Rodrigues-Silva, J., Silva-Hormaz bal, M., & Alsina,  . (2023). Concepciones del alumnado sobre ingenier a y sus conexiones con las matem ticas y las ciencias. *Ense anza de las Ciencias. Revista de Investigaci n y Experiencias Did cticas*, 41(3), 33–51. [CrossRef]
- Roehrig, G. H., Dare, E. A., Ring-Whalen, E., & Wieselmann, J. R. (2021). Understanding coherence and integration in integrated STEM curriculum. *International Journal of STEM Education*, 8(1), 2. [CrossRef]
- Saxton, E., Burns, R., Holveck, S., Kelley, S., Prince, D., Rigelman, N., & Skinner, E. A. (2014). A common measurement system for K-12 STEM education: Adopting an educational evaluation methodology that elevates theoretical foundations and systems thinking. *Studies in Educational Evaluation*, 40, 18–35. [CrossRef]
- Shernoff, D., Sinha, S., Bressler, D., & Ginsburg, L. (2017). Assessing teacher education and professional development needs for the implementation of integrated approaches to STEM education. *International Journal of STEM Education*, 4, 13. [CrossRef]
- Tan, A.-L., Ong, Y. S., Ng, Y. S., & Tan, J. H. J. (2023). STEM problem solving: Inquiry, concepts, and reasoning. *Science & Education*, 32(2), 381–397. [CrossRef]
- Tao, Y. (2019). Kindergarten teachers’ attitudes toward and confidence for integrated STEM education. *Journal for STEM Education Research*, 2(2), 154–171. [CrossRef]
- Thibaut, L., Ceuppens, S., Loof, H. D., Meester, J. D., Goovaerts, L., Struyf, A., Pauw, J. B., Dehaene, W., Deprez, J., Cock, M. D., Hellinckx, L., Knipprath, H., Langie, G., Struyven, K., de Velde, D. V., Petegem, P. V., & Depaepe, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education*, 3(1), 2. [CrossRef] [PubMed]
- Thibaut, L., Knipprath, H., Dehaene, W., & Depaepe, F. (2019). Teachers’ attitudes toward teaching integrated STEM: The impact of personal background characteristics and school context. *International Journal of Science and Mathematics Education*, 17(5), 987–1007. [CrossRef]
- Weinberg, A. E., Balgopal, M. M., & Sample McMeeking, L. B. (2021). Professional growth and identity development of STEM teacher educators in a community of practice. *International Journal of Science and Mathematics Education*, 19(S1), 99–120. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.