Journal of Education and Learning (EduLearn)
Vol. 20, No. 4, November 2026, pp. 2047~2058
ISSN: 2089-9823, DOI: 10.11591/edulearn.v20i4.25392

The impact of scientific inquiry on teacher self-efficacy in the Chilean context

Katherine Acosta-García[1], Marlene Morales Choque[2], Juan Pablo Aguirre Gonzalez[1], Sandra Ramírez Hernández[3], Sara Maldonado-Araya[3]
[1]Department of Education, Faculty of Education and Humanities, University of Tarapacá, Arica, Chile
[2]Department of Chemistry, Faculty of Sciences, University of Tarapacá, Arica, Chile
[3]Department of Education, Doctoral Program in Education, University of Tarapacá, Arica, Chile

Article Info

Article history:

Received Jun 25, 2025
Revised Apr 2, 2026
Accepted Jun 8. 2026

Keywords:

Continuous professional development
In-service teachers
Science education
Scientific inquiry
Teacher self-efficacy

ABSTRACT

Teacher self-efficacy directly influences the quality of science teaching and student learning. This study examined the impact of an inquiry-based continuous professional development program on 94 public school science teachers from northern Chile, employing a pre-experimental design with pre- and post-test measurements administered through the Ohio State Teacher Efficacy Scale (OSTES). The results showed statistically significant improvements in instructional strategies (d=1.19), classroom management (d=1.02), and student engagement (d=1.31), with increases ranging from 11.7% to 19.2%. These findings suggest that continuous professional development focused on inquiry-based pedagogical approaches can substantially strengthen teachers' self-efficacy beliefs and enhance science education. Nevertheless, the absence of a control group in the pre-experimental design limits causal inference.

Corresponding Author:

Katherine Acosta-García
Department of Education, Faculty of Education and Humanities, University of Tarapacá
18 de Septiembre #2222, Arica, Chile
Email: kacostag@academicos.uta.cl

1. INTRODUCTION

High-quality science education faces a fundamental challenge involving the effective implementation of innovative pedagogical methodologies that promote meaningful learning. Within this context, scientific inquiry has emerged as a key pedagogical approach that transforms the classroom into a space where students actively construct knowledge and develop higher-order cognitive skills, with documented positive effects on their engagement with and attitudes toward science [1]. By emphasizing exploration, reasoning and problem-solving, this instructional model enhances conceptual understanding of natural phenomena and fosters self-efficacy in students themselves [2]. Its relevance lies in the global need to develop scientific literacy as a core competency for informed civic participation, an internationally recognized priority in contemporary science education.

However, international research reveals significant shortcomings in the implementation of inquiry-based science education across all educational levels. A study encompassing 32 countries found that students completing secondary education often demonstrate inadequate understanding of scientific inquiry processes, a pattern replicated from primary through high school levels [3]. This issue is largely rooted in teachers' limited competencies to effectively apply inquiry-based pedagogical approaches, resulting in a vicious cycle in which traditional teaching methods perpetuate suboptimal learning outcomes in science.

The effective implementation of scientific inquiry requires that teachers possess not only solid disciplinary and methodological knowledge, but also a high level of teacher self-efficacy. Bandura [4] defined this construct as the belief in one's own capability to organize and execute the actions required to manage prospective situations. In the educational context, teacher self-efficacy refers to educators' beliefs about their ability to positively influence students' learning and behavior, even in challenging environments [5]. The literature has consistently shown that teachers with higher levels of self-efficacy are more likely to implement innovative strategies, maintain high expectations for student achievement, and demonstrate greater persistence when facing educational challenges [6]–[8].

Within this framework, teachers' confidence in their ability to guide inquiry-based processes is crucial for overcoming the inherent challenges of this methodology, including managing uncertainty, responding to students' unpredictable questions, and integrating cognitively activating strategies [6]. Research conducted in the Chilean context has shown that science teachers exhibit varying levels of self-efficacy across different dimensions of their teaching practice, with instructional strategies identified as the area most in need of support [9]. At the same time, evidence suggests that teachers with higher self-efficacy in science education are more likely to implement inquiry-based practices, as they feel better equipped to overcome obstacles and persist in refining their pedagogical approaches [10]. The development of teacher self-efficacy is fostered through continuous professional development (CPD), prior teaching experience, and institutional environments that support innovation and peer collaboration [11].

To address this issue, the Chilean Ministry of Education implemented the Scientific Inquiry for Science Education Program (ICEC, its acronym in Spanish), an initiative aimed at improving science teaching and learning through scientific inquiry and teacher professional development [12]. Likewise, evidence from the ICEC context indicates that participation in the program is perceived by teachers as a significant professional development opportunity, associated with increased confidence in science teaching, the incorporation of new methodological strategies, and the application of pedagogical tools to improve classroom practice [13]. Furthermore, it incorporates blended learning (b-learning) modalities, which have shown significant positive effects on teachers' achievement and self-efficacy outcomes by combining in-person and online instructional components [14].

Despite the theoretical recognition of the relevance of teacher self-efficacy, a significant gap remains in the empirical evidence on how long-term inquiry-based CPD programs influence this psychological construct. This gap is particularly evident in Latin America, where educational contexts present specific challenges related to high cultural diversity and socioeconomic vulnerability.

Based on the above, this study asks about to what extent can a CPD program based on scientific inquiry modify self-efficacy beliefs in natural science teachers working in public schools in northern Chile? Accordingly, the general objective is to analyze the impact of a scientific inquiry-based CPD program on teachers' self-efficacy beliefs through the implementation of the ICEC program. The geographic context adds value by focusing on northern Chile, characterized by ethnocultural diversity and the presence of indigenous communities [15], as well as by student vulnerability and shortages of teachers in science, technology, engineering and mathematics or STEM areas [16], offering evidence on the implementation of educational policy in diverse and underserved settings.

To achieve this general aim the research addresses the specific dimensions of teacher self-efficacy through the following specific objectives: i) to evaluate changes in self-efficacy beliefs related to instructional strategies among natural science teachers following participation in the scientific inquiry-based CPD program; ii) to determine modifications in self-efficacy beliefs concerning classroom management among natural science teachers after the implementation of the inquiry-based CPD program; iii) to analyze transformations in self-efficacy beliefs regarding student engagement among natural science teachers, through the application of the CPD program; and iv) to compare the magnitude of the program's effects on the three dimensions of teacher self-efficacy, across different teacher cohorts in order to identify effectiveness patterns and inform future program optimizations.

2. METHOD

2.1. Research design

This study employed a longitudinal pre-experimental design with repeated measures (pre-test/post-test), specifically classified as a single-group O1-X-O2 design, in accordance with methodological approaches recommended for assessing the initial impact of educational interventions [17], [18]. From a quantitative perspective, the study enabled the collection of objective and comparable measurements of the construct of teacher self-efficacy, which has been shown to be amenable to empirical analysis through validated instruments [19]. Moreover, this approach supports the application of statistical tests that allow for the identification of generalizable trends within large and contextually grounded samples [20]. Nevertheless, the threats to internal validity inherent in this design are acknowledged, particularly maturation effects and the influences of repeated assessment processes, which limit the ability to establish definitive causal inferences.

2.2. Participants and context

The sample consisted of 94 in-service natural science teachers selected through convenience sampling [21], [22]. The demographic characteristics of the sample include a mean age of 42.3 years (SD=8.7) and an average teaching experience of 12.5 years (SD=6.2). Inclusion criteria required a minimum of two years of professional teaching experience. Additionally, only teachers from public schools were included in order to maintain sample homogeneity regarding school management type. Of these schools, 12% were rural and the remaining were urban.

The participants were organized into three consecutive cohorts of 31, 30 and 33 teachers respectively. The different cohorts were carried out over eight consecutive months each. This cohort-based organization corresponded to the staggered implementation of the professional development program. The study was conducted in four municipalities in the Arica and Parinacota region, Chile, as shown in Table 1, an area characterized by the geographical dispersion of schools, the sociocultural diversity of the student population and vulnerability levels above the national average [23].

Table 1. Demographic characteristics of participants (N=94)

Characteristic	Category	n	%
Gender	Women	62	66.0
	Men	32	34.0
School location	Urban	83	88.0
	Rural	11	12.0
Teaching experience (years)	2-4	24	25.5
	5-9	34	36.2
	10+	36	38.3

Note: percentages computed over non-missing data; totals may not equal 100% due to rounding

2.3. Instrument and implementation

The instrument used for data collection was the Ohio State Teacher Efficacy Scale (OSTES), selected due to its internationally recognized psychometric validity and its appropriateness for assessing teacher self-efficacy in educational contexts focused on science instruction. The questionnaire was developed by Tschannen-Moran and Hoy [5] and has demonstrated high reliability across various cross-cultural applications, with reported Cronbach's alpha coefficients ranging from 0.87 to 0.95 across its different dimensions [24], [25]. The instrument corresponds to the full 24-item version of the OSTES [5]. Items are assessed using a 9-point scale with five verbal anchors: 1 (nothing), 3 (very little), 5 (some influence), 7 (quite a bit), and 9 (a great deal), allowing intermediate responses between these points. For this study, the instrument was translated into Spanish through a translation-back-translation process conducted by bilingual specialists in science education.

The first dimension, efficacy in instructional strategies, assesses teachers' confidence in selecting, adapting and implementing appropriate pedagogical methods. This dimension is particularly relevant in inquiry-based science teaching contexts where planning and facilitating open-ended and student-centered learning experiences entails a high level of complexity. The second dimension, efficacy in classroom management refers to teachers perceived ability to establish and maintain structured and supportive learning environments. This competence is especially critical when working with active methodologies such as inquiry-based learning which require simultaneous management of diverse group dynamics and physical resources. The third dimension, efficacy in student engagement focuses on teachers perceived ability to motivate students and foster their active and sustained participation. This dimension is considered essential to the success of inquiry-based approaches, as it promotes student autonomy, curiosity and deep engagement in the knowledge construction process.

To measure each of these dimensions, the instrument includes specific items. For example: i) efficacy for instructional strategies: "To what extent can you use a variety of assessment strategies?" (item 1), "How well can you respond to difficult questions from your students?" (item 5); ii) efficacy for classroom management: "How much can you do to calm a student who is disruptive or noisy?" (item 11), "How well can you respond to defiant students?" (item 14); and iii) efficacy for student engagement: "How much can you do to get students to believe they can do well in schoolwork?" (item 17), "How much can you do to get through to the most difficult students?" (item 24).

The OSTES questionnaire was administered following a standardized protocol. The pre-test was conducted during the first week of the CPD program prior to any instruction related to scientific inquiry. The post-test was administered at the end of the course with an approximate interval of eight months between the two measurements, thereby ensuring the temporal comparability of the data.

2.4. Data analysis

The statistical analysis was structured into three progressive phases, following methodological guidelines recommended for intervention research in applied educational contexts [26]. The first phase involved descriptive analyses to characterize the distribution of pre-test and post-test scores. Measures of central tendency (mean, median), dispersion (standard deviation) and distribution shape (skewness and kurtosis) were calculated to identify potential deviations from the parametric assumptions required for subsequent inferential procedures. The internal consistency of the OSTES dimensions was assessed using Cronbach's alpha coefficient, with values equal to or greater than α=0.70 considered indicative of acceptable reliability.

In the second phase, assumptions of normality and homogeneity were assessed. The Shapiro-Wilk test, recommended for samples with fewer than 50 cases [27] was used. A significance level of α=0.05 was established for all statistical tests. The test was useful to determine whether the distributions significantly deviated from normality. In addition, homogeneity of variance between pre-test and post-test measurements was examined to confirm the stability of score dispersion.

The third phase focused on the application of paired sample t-tests to assess significant differences between pre-test and post-test scores across the three dimensions of the OSTES questionnaire. This inferential technique was selected based on the repeated-measures design of the study and the prior confirmation of statistical assumptions [26]. Given that the analyses were conducted on theoretically independent dimensions of the teacher self-efficacy construct, no alpha correction for multiple comparisons was applied. Additionally, effect size was estimated using Cohen's d, with 95% confidence intervals calculated using the bootstrapping method with 1000, iterations, in order to determine the practical magnitude of the observed changes. Standard interpretative thresholds were used: small (d=0.20), medium (d=0.50) and large (d=0.80) [27]. All statistical analyses were conducted using IBM SPSS Statistics software, version 26.

2.5. Intervention

The CPD program under evaluation was structured into eight thematic units, with a total duration of 400 academic hours distributed as 60 face-to-face activities and 40 online components. The program was delivered by a team of five specialized trainers, all holding a master's or doctoral degree in education and with a minimum of 10 years of experience in teacher training. To ensure implementation fidelity, a standardized protocol was used that included: i) detailed instructional guides for each session; ii) weekly coordination meetings among the training team; and iii) cross-observation among trainers in at least 20% of the face-to-face sessions. This blended b-learning structure allowed for a balance between theoretical deepening and practical experiences facilitating both pedagogical reflection and the contextualized application of content.

The first unit, conceptions of science, addressed the epistemological foundations of science, emphasizing its dynamic empirical nature and its embeddedness in social and cultural contexts. The second unit, conceptions of science education, focused on scientific literacy as a key goal of science teaching highlighting its role in fostering active and informed citizenship. Subsequent units addressed critical areas such as the analysis of the national curriculum, the implementation of scientific inquiry in the classroom, assessment for learning, the design of teaching resources, practitioner research and the development of teacher learning communities. This training sequence was designed with a progressive and integrated approach, aligned with the competencies required for effective inquiry-based science teaching. A detailed breakdown of each unit including objectives and allocated hours is presented in Table 2.

2.6. Ethical considerations

This study was conducted in accordance with ethical standards for educational research involving human participants and was approved by the Institutional Ethics Committee of the University of Tarapacá, an accredited institution. Prior to data collection, ethical approval was obtained from the Institutional Ethics Committee of University of Tarapacá. All participants provided written informed consent after receiving detailed information about the study's objectives, procedures, voluntary nature of participation, and data confidentiality measures. The informed consent process emphasized that participation was entirely voluntary and that teachers could withdraw from the study at any time without any professional consequences. Additionally, all data collection and analysis procedures adhered to strict confidentiality protocols, ensuring that individual responses could not be linked to specific participants in any reports or publications. The study design and implementation followed the ethical guidelines established by the Chilean Ministry of Education for research conducted in public educational institutions.

Table 2. Structure of the specialization course in scientific inquiry: units, objectives and hour allocation

Unit	HP	HO	HT	Objective
Conceptions of science	12	10	22	To understand and value science as the result of a human activity, carried out through diverse procedures, forming an integrated and interdependent body of knowledge methods and attitudes. It is empirical, dynamic, subject to change in light of new evidence and influenced by the political, social and cultural contexts in which it develops.
Conceptions of science education	12	10	22	To understand the importance of science teaching and learning in schools. As it contributes to students' scientific literacy, enabling them to actively and responsibly participate in evidence-based decision-making that affects personal wellbeing, their community and the environment.
Science curriculum	16	8	24	To deepen understanding of the current curriculum, recognizing and valuing its aims, objectives, rationale and curricular progression as a tool for achieving learning outcomes in natural sciences.
Scientific inquiry in the classroom	80	30	110	To recognize scientific inquiry as an essential component of scientific practice and as a pedagogical approach that enables students to develop and understand scientific ideas by engaging in the processes of doing science.
Assessment for learning	34	18	52	To appropriate the assessment for learning model in order to define and develop various assessment strategies in natural sciences, aligned with scientific inquiry.
Design of educational resources	24	20	44	To design, develop, evaluate and/or adapt various educational resources to support the implementation of scientific inquiry in natural science classrooms considering the local context.
Classroom-based research	34	24	58	To apply methodological tools to examine and investigate science teaching practices, promoting individual and collaborative reflection aimed at continuous improvement of science teaching and learning processes.
Learning communities	28	40	68	To understand learning communities as a model of collective reflection and professional learning that values educator's knowledge and experiences within a collaborative peer-based framework.
Total course hours	240	160	400	Percentage distribution: 60% face-to-face hours, 40% on-line hours.

Note: HP=face-to-face hours, HO=on-line hours, and HT=total hours.

3. RESULTS AND DISCUSSION

3.1. Instrument reliability

The internal consistency of the OSTES questionnaire was assessed using Cronbach's alpha coefficient for each cohort, based on both pre-test and post-test measurements. The results indicate satisfactory reliability values ranging from 0.741 to 0.930, supporting the instrument's adequacy for measuring teacher self-efficacy in this educational context, as shown in Table 3. According to widely accepted psychometric standards, an alpha value of ≥0.70 is considered acceptable for establishing internal reliability in instruments undergoing validation [28], [29].

Table 3. Reliability analysis results for the OSTES instrument

Cohort	Pre-test	Post-test	Items
1	0.741	0.909	24
2	0.770	0.930	24
3	0.752	0.917	24

3.2. Descriptive changes in teacher self-efficacy

Descriptive analysis of the scores obtained on the OSTES instrument revealed positive changes across all three dimensions of teacher self-efficacy following the implementation of the CPD program. In every case, post-test scores consistently exceeded pre-test values, suggesting a favorable effect of the training intervention. Table 4 presents a summary of the means, standard deviations, and percentage gains for each dimension. In the instructional strategies dimension, there was an average increase of 1.2 points (SD=0.8), rising from a mean of 7.6 in the pre-test to 8.8 in the post-test, representing a relative improvement of 15.8%. This outcome exceeds the typical gains reported in recent meta-analyses on the effectiveness of training interventions focused on pedagogical strategies in science education [30].

Moreover, research on teacher professional trajectories highlights that such gains are not only attributable to the program's thematic content, but also to its reflective and progressive structure, which enables teachers to reinterpret their practice through the contextualized application of active methodologies [31].

Regarding the classroom management dimension, scores increased by an average of 0.9 points (SD=0.7), reflecting a relative gain of 11.7%, from 7.7 to 8.6 points. While this change was positive, it was less pronounced compared to the other dimensions. In this regard, the literature suggests that improvement in this area often requires longer-term interventions and systematic follow-up to translate into sustainable transformations, as classroom management skills are typically more deeply rooted in teachers' prior experience and specific contextual factors [32].

Finally, the student engagement dimension showed the most significant gain, with an average increase of 1.5 points (SD=0.9), equivalent to a relative improvement of 19.2%, from 7.8 to 9.3 points. This finding aligns with studies that have demonstrated the positive impact of programs integrating personalized learning, educational technologies, and student-centered pedagogical strategies factors that have been associated with a substantial strengthening of teachers' beliefs in their ability to build meaningful connections with students and foster active classroom participation [33]. The results are summarized in Figure 1, which displays a comparative overview of pre- and post-test means by OSTES dimension, along with their respective standard deviations.

Table 4. Descriptive statistics and percentage gain by OSTES dimension

Dimension	Pre-test M (DE)	Post-test M (DE)	Gain (points)	Gain (%)
Instructional strategies	7.6 (0.8)	8.8 (0.7)	1.2	15.8%
Classroom management	7.7 (0.9)	8.6 (0.8)	0.9	11.7%
Student engagement	7.8 (1.0)	9.3 (0.7)	1.5	19.2%

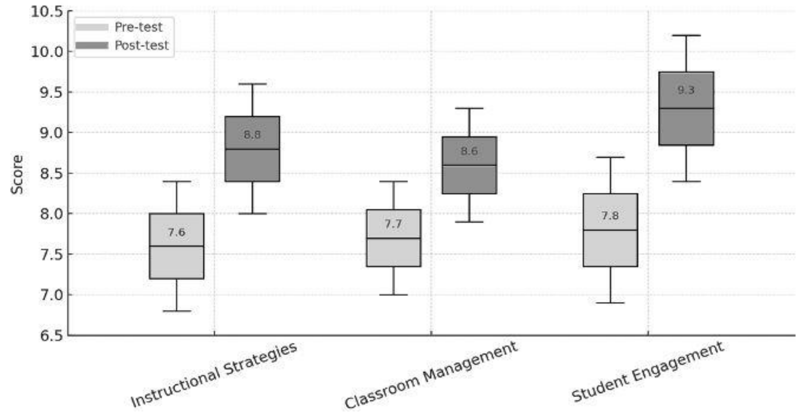

Figure 1. Distribution of pre- and post-test scores by OSTES dimension (values centered on the mean ± standard deviation)

3.3. Inferential analysis and effect size

Paired-sample t-tests confirmed statistically significant differences between pre-test and post-test scores across all three dimensions of teacher self-efficacy. The instructional strategies dimension results showed $t(93)=11.54$, $p<0.001$, with a large effect size ($d=1.19$). For the classroom management dimension, the result was $t(93)=9.87$, $p<0.001$, also indicating a large effect ($d=1.02$). In the student engagement dimension, the strongest effect was observed: $t(93)=12.73$, $p<0.001$, with the largest observed effect size ($d=1.31$). These results are summarized in Table 5.

Table 5. Results of t-tests and effect sizes by self-efficacy dimension

Dimension	t(93)	p-value	d (Cohen)	CI 95% for d
Instructional strategies	11.54	<0.001	1.19	[0.95, 1.43]
Classroom management	9.87	<0.001	1.02	[0.79, 1.25]
Student engagement	12.73	<0.001	1.31	[1.06, 1.56]

These effect sizes are considered very large according to Cohen's reference parameters, where values of $d \geq 0.80$ are classified as large effects. These effect sizes substantially exceed those reported in recent meta-analyses on the effectiveness of teacher professional development programs focused on self-efficacy.

Contemporary research indicates that well-designed training interventions typically yield moderate to high effects, most commonly within the range of d=0.60 to 0.90 [32]. The comparatively larger effect sizes observed in this study (d=1.19 to 1.31) may be attributed to multiple factors. First, the extended duration of the program (400 hours) allowed for a systematic approach to the four sources of self-efficacy proposed by Bandura [4]: i) mastery experiences, through the practical implementation of inquiry-based teaching in real classroom contexts; ii) vicarious experiences, via the observation of peers modeling successful strategies; iii) social persuasion, facilitated by continuous mentoring from specialized trainers and peer feedback; and iv) positive physiological and emotional states, fostered by a collaborative professional learning environment. This deliberate integration of all four sources contrasts with shorter programs that typically emphasize only one or two of these dimensions [34]. Consistent with this theoretical framework, previous studies have indicated that higher levels of teacher self-efficacy are also associated with more positive emotional experiences in the classroom, such as greater enjoyment of teaching and lower levels of anxiety, reinforcing the role of emotional regulation as a relevant component of teachers' professional performance [35].

Nevertheless, it is important to note that not all studies report effects of this magnitude. For example, Liu *et al*. [36] found moderate effects (g=0.55) in a meta-analysis of professional development programs for STEM science teachers, while Zhou *et al*. [37] reported an effect size of g=0.64 in programs for K-12 science teachers. These differences may be explained by: i) the specific emphasis on systematic reflection on teaching practice in our program, an element that may be absent or less pronounced in shorter interventions; ii) the modular structure that enabled iterative cycles of application-reflection-refinement over 18 months; and iii) the context of high educational vulnerability among participants, where professional development needs may be more pronounced, generating greater potential for improvement from lower baseline levels.

Additionally, recent studies emphasize that the integration of educational technologies into professional development programs enhances teachers' self-efficacy beliefs and confidence in addressing complex instructional challenges [38]. This effect is further amplified when teacher well-being is considered a structural component of professional development, given its documented impact on teaching practice and professional outcome [39].

3.4. Comparative post-test analysis by cohort

The comparative analysis among the three cohorts revealed consistent patterns of improvement across all dimensions of teacher self-efficacy in the post-test, although with slight variations in the magnitude of effects. Cohort 1 showed the most pronounced gains in each OSTES dimension (instructional strategies: d=1.34; classroom management: d=1.18; and student engagement: d=1.47), while cohorts 2 and 3 exhibited similarly high but slightly lower effect sizes (cohort 2: d=1.12, 0.94, 1.23; cohort 3: d=1.09, 0.95, 1.22, respectively). The exact values of the effect sizes by cohort are presented in Table 6.

Despite these numerical differences, the analysis of variance revealed no statistically significant differences between cohorts (F(2, 91)=1.87, p=0.16), suggesting a homogeneous level of program effectiveness across groups. A plausible explanation for Cohort 1's advantage could lie in an organizational learning effect on the part of the program facilitators, who may have optimized their training strategies based on insights gained from the initial implementation. Alternatively, the variations may be attributable to characteristics specific to the participants, such as prior teaching experience, initial teacher education, or the institutional contexts in which they work. Future studies should incorporate multivariate analyses to control for these potential confounding variables and to deepen the interpretation of the observed effects.

Table 6. Effect sizes (Cohen's d) by cohort and OSTES dimension

Cohort	Instructional strategies (d)	Classroom management (d)	Student engagement (d)
Cohort 1 (n=31)	1.34	1.18	1.47
Cohort 2 (n=30)	1.12	0.94	1.23
Cohort 3 (n=33)	1.09	0.95	1.22
Global (N=94)	1.19	1.02	1.31

3.5. Non-parametric analysis using the Wilcoxon test

Given that some individual OSTES items did not meet the assumptions of normality according to the Shapiro-Wilk test (p<0.05), particularly in the distribution of pre-test/post-test differences, the inferential analysis was complemented using the nonparametric Wilcoxon signed-rank test for each item individually. The use of the Wilcoxon signed-rank test revealed statistically significant transformations (p<0.05) in 11 of the 24 items of the OSTES instrument when comparing pre-test and post-test responses. A notable aspect of these findings lies in the fact that all 11 items demonstrated consistent patterns of statistical significance across the three analyzed cohorts, providing methodological robustness and confirming the replicability of the observed effects.

The pattern of changes reflects a favorable shift in teachers' self-efficacy beliefs, primarily concentrated in two core constructs of the instrument: the efficacy in instructional strategies dimension and the efficacy in student engagement dimension. Table 7 details the specific items that exhibited significant change, organized according to these conceptual categories. This systematization facilitates the identification of particular aspects of teaching practice that experienced the most pronounced transformations as a result of the CPD intervention.

Table 7. Items showing statistically significant change according to dimensions by OSTES

Dimensions	Statement items
Efficacy in instructional strategies	01. To what extent can you use a variety of assessment strategies?
	02. To what extent can you provide an alternative explanation or example when students are confused?
	03. How well can you craft good questions for your students?
	04. How well can you implement alternative teaching strategies in your classroom?
	05. How well can you respond to your students' difficult questions?
	06. How much can you do to tailor your lessons to the appropriate level for individual students?
	08. How well can you provide appropriate challenges for very capable students?
Efficacy in student engagement	18. How much can you do to help your students to value learning?
	19. How much can you do to motivate students who show low interest in schoolwork?
	22. How much can you do to help your students think critically?
	23. How much can you do to foster student creativity?

3.6. Subgroup analyses

We explored whether pre-post change in teacher self-efficacy differed by gender and by years of teaching experience. Change scores were computed as Δ=post-pre for each OSTES dimension. Gender: an independent-samples t-test compared Δ between women and men. Differences were not statistically significant across dimensions (all $p>0.05$). Robust non-parametric Mann–Whitney tests yielded comparable conclusions. The effect size for gender comparisons was small ($r<0.10$), indicating differences of trivial magnitude between groups. Teaching experience: participants were grouped into 2-4, 5-9, and 10+ years of experience. One-way ANOVAs on Δ showed no significant differences among experience groups (all $p>0.05$). Where normality assumptions were questionable, Kruskal-Wallis tests led to the same inference. The effect size was small ($\eta^2<0.01$), suggesting that teaching experience explains a minimal proportion of the variability in the observed changes. Assumptions: normality (Shapiro-Wilk) and homogeneity (Levene) were examined before parametric tests; when violated, we report robust alternatives.

3.7. Limitation and future research

This pre-experimental study has several limitations: i) self-selection bias: participation was voluntary, so teachers who enrolled may differ from non-participants in motivation or openness to innovation; ii) lack of observational triangulation: outcomes rely on self-report; systematic classroom observations and student outcomes were not collected; iii) design constraints: without a control group, alternative explanations (history, maturation, and testing effects) cannot be ruled out; findings should be interpreted as associative rather than causal; and iv) regional specificity: the sample draws from public schools in northern Chile, which may limit generalizability.

Future studies should incorporate comparison groups (e.g., wait-list controls), multi-method measures (observations, student assessments, and artifacts), longer follow-ups to assess durability, and sampling across diverse regions and school types. Subgroup analyses can be powered a priori to examine equity-relevant moderators (e.g., school socioeconomic status, rurality, and contract type). Likewise, future research could examine whether strengthening teacher self-efficacy facilitates the adoption of more complex pedagogical approaches, such as the use of socioscientific issues, which can support the development of critical thinking in teacher education [40]. Consistent with the study's pre-experimental design, the results should be interpreted as associative; the data suggest improvements in self-efficacy following participation in the CPD program.

4. CONCLUSION

This study demonstrated that a CPD program focused on scientific inquiry can lead to significant improvements in teacher self-efficacy, particularly in instructional strategies, classroom management, and student engagement. The consistency of results across cohorts and the magnitude of observed effects confirm that well-structured, long-term interventions grounded in robust theoretical foundations can positively transform science teachers' professional beliefs.

Wilcoxon test analyses revealed that the most substantial changes were concentrated in instructional strategies and student engagement, with statistically significant improvements in 11 items. The most impactful elements include diversification of assessment strategies, provision of alternative explanations, formulation of effective questions, implementation of differentiated instruction, and skills related to motivating disengaged students, fostering critical thinking, and stimulating creativity.

These findings support the inclusion of long-term inquiry-based CPD (≥400 hours) as a national priority to strengthen science education quality, particularly in regions with high educational vulnerability. Educational authorities should consider sustained funding for modular programs ensuring practical classroom implementation, peer observation, mentoring, and collaborative reflection. In teacher education, results indicate that professional development programs should prioritize designs that integrate theory and practice through iterative cycles of application–reflection–refinement over extended periods, with particular emphasis on strengthening self-efficacy related to student engagement.

Future research should address this study's limitations through controlled experimental designs with random assignment. Additionally, longitudinal studies assessing teacher self-efficacy 12- and 24-months post-intervention would be valuable, examining relationships with observable changes in classroom practices and student learning outcomes.

FUNDING INFORMATION

This research was funded by the Agencia Nacional de Investigación y Desarrollo (ANID) (the National Research and Development Agency of Chile) through the Fondecyt project 11241075 titled: Teaching Natural Sciences at the Primary Level in a Border Context: Relationships and Tensions Between Professional Practice, Initial Teacher Training, and Self-Efficacy.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Katherine Acosta-García	✓	✓	✓	✓		✓	✓		✓	✓		✓	✓	✓
Marlene Morales Choque		✓				✓		✓	✓			✓		
Juan Pablo Aguirre González	✓		✓	✓		✓			✓			✓	✓	
Sandra Ramírez Hernández		✓		✓	✓			✓		✓	✓			
Sara Maldonado-Araya	✓		✓		✓			✓		✓	✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

The research related to human use has been complied with all the relevant national regulations and institutional policies in accordance with the tenets of the Helsinki Declaration and has been approved by the authors' institutional review board or equivalent committee.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [KA-G], upon reasonable request.

REFERENCES

[1] R. P. Antonio and M. S. Prudente, "Effects of inquiry-based approaches on students' higher-order thinking skills in science: a meta-analysis," *International Journal of Education in Mathematics, Science and Technology*, vol. 12, no. 1, pp. 251–281, 2024, doi: 10.46328/ijemst.3216.

[2] C. W. Beck and L. S. Blumer, "The relationship between perceptions of instructional practices and student self-efficacy in guided-inquiry laboratory courses," *CBE—Life Sciences Education*, vol. 20, no. 1, p. ar8, Mar. 2021, doi: 10.1187/cbe.20-04-0076.

[3] J. S. Lederman *et al.*, "Completing the progression establishing an international baseline of primary, middle and secondary students' views of scientific inquiry," *International Journal of Science Education*, vol. 46, no. 7, pp. 715–731, Nov. 2023, doi: 10.1080/09500693.2023.2256458.

[4] A. Bandura, *Self-efficacy: the exercise of control*. 1st ed, New York: W.H. Freeman and Company, 1997.

[5] M. Tschannen-Moran and A. W. Hoy, "Teacher efficacy: capturing an elusive construct," *Teaching and Teacher Education*, vol. 17, no. 7, pp. 783–805, Oct. 2001, doi: 10.1016/s0742-051x(01)00036-1.

[6] M. Ryan and G. D. Hendry, "Sources of teacher efficacy in teaching reading: success, sharing, and support," *The Australian Journal of Language and Literacy*, vol. 46, no. 1, pp. 1–14, Dec. 2022, doi: 10.1007/s44020-022-00016-0.

[7] Ş. Orakcı, D. Yüreğilli Göksu, and S. Karagöz, "A mixed methods study of the teachers' self-efficacy views and their ability to improve self-efficacy beliefs during teaching," *Frontiers in Psychology*, vol. 13, Jan. 2023, doi: 10.3389/fpsyg.2022.1035829.

[8] S. Alshuhumi *et al.*, "Examining the impact of Omani primary school climate and teacher self-efficacy on innovative teaching practices: a structural equation modeling approach," *Frontiers in Education*, vol. 10, Mar. 2025, doi: 10.3389/feduc.2025.1487857.

[9] K. Acosta-García and M. Navarro-Ibañez, "Teacher self-efficacy in Northern Chile: challenges in instructional strategies," *International Journal of Learning, Teaching and Educational Research*, vol. 24, no. 1, pp. 315–334, Jan. 2025, doi: 10.26803/ijlter.24.1.16.

[10] V. D. E. Welter, M. Dawborn-Gundlach, L. Großmann, and M. Krell, "Adapting a self-efficacy scale to the task of teaching scientific reasoning: collecting evidence for its psychometric quality using Rasch measurement," *Frontiers in Psychology*, vol. 15, Feb. 2024, doi: 10.3389/fpsyg.2024.1339615.

[11] A. A. Christensen and J. Jerrim, "Professional learning communities and teacher outcomes. a cross-national analysis," *Teaching and Teacher Education*, vol. 156, p. 104920, Apr. 2025, doi: 10.1016/j.tate.2024.104920.

[12] E. Hernández Lémann and D. Caffi Pizarro, "The scientific research program for science education," (in Spanish), *Revista de Innovación en Enseñanza de las Ciencias*, vol. 4, no. 1, pp. 3–12, Aug. 2020, doi: 10.5027/reinnec.V4.I1.69.

[13] L. V. Castro, A. Maldonado-Márquez, K. G. Beltrán, and P. A. Ojeda, "Teacher perceptions of the implementation of the inquiry-based approach in science teaching in b-learning mode: perspectives from Magallanes, Chile," (in Spanish), *Digital Education Review*, no. 47, pp. 331–346, 2025, doi: 10.1344/der.2025.47.331-346.

[14] R. F. Schmid, E. Borokhovski, R. M. Bernard, D. I. Pickup, and P. C. Abrami, "A meta-analysis of online learning, blended learning, the flipped classroom and classroom instruction for pre-service and in-service teachers," *Computers and Education Open*, vol. 5, p. 100142, Dec. 2023, doi: 10.1016/j.caeo.2023.100142.

[15] V. Rivera and A. Paulsen, "Education as a homogenizing engine in indigenous cultures: case study of Aymara children in Northern Chile," *Discover Education*, vol. 2, art. no. 8, 2023, doi: 10.1007/s44217-023-00030-5.

[16] K. Acosta-García, E. Valdivia, J. Jiménez, M. Dueñas-Zorrilla, C. Mondaca, and C. Alfaro-Contreras, "Situated science education and curricular justice in rural borderland schools: elementary teachers' voices from Northern Chile," *Education Sciences*, vol. 15, no. 12, article 1656, 2025, doi: 10.3390/educsci15121656.

[17] S. Em, "Exploring experimental research: methodologies, designs, and applications across disciplines," *SSRN Electronic Journal*, 2024, doi: 10.2139/ssrn.4801767.

[18] Teachers Institute, "Pre-experimental designs in educational research: an overview," 2025.

[19] S. Woodcock, N. Tournaki, and J. Ehrich, "Teacher self-efficacy: validating a new measurement scale to capture the elusive construct," *British Educational Research Journal*, vol. 51, no. 5, pp. 2229–2255, Apr. 2025, doi: 10.1002/berj.4163.

[20] J. Täschner, T. Dicke, S. Reinhold, and D. Holzberger, "Yes, I can!' A systematic review and meta-analysis of intervention studies promoting teacher self-efficacy," *Review of Educational Research*, vol. 95, no. 1, pp. 3–52, Jan. 2024, doi: 10.3102/00346543231221499.

[21] K. K. Harrall *et al.*, "Power and sample size analysis for longitudinal mixed models of health in populations exposed to environmental contaminants: a tutorial," *BMC Medical Research Methodology*, vol. 23, no. 1, Jan. 2023, doi: 10.1186/s12874-022-01819-y.

[22] H. Al-Nashash *et al.*, "Computation of statistical power and sample size for in vivo research models," *arXiv:2505.19666*, May 2025, doi: 10.48550/arXiv.2505.19666.

[23] P. Zapata-Sepúlveda, N. B. Villalobos, P. M. Espinoza-Concha, and V. Beserra, "Intercultural education: an analysis of the print media in Arica (Chile) and Tacna (Peru)," *Convergencia*, vol. 32, e24544, 2025, doi: 10.29101/crcs.v32i1.24544.

[24] F. Salas-Rodríguez, S. Lara, and M. Martínez, "Spanish version of the teachers' sense of efficacy scale: an adaptation and validation study," *Frontiers in Psychology*, vol. 12, p. 714145, 2021, doi: 10.3389/fpsyg.2021.714145.

[25] L. Gálvez-Nieto, S. Salvo-Garrido, S. Domínguez-Lara, K. Polanco-Levicán, and M. Mieres-Chacaltana, "Psychometric properties of the teachers' sense of efficacy scale in a sample of Chilean public school teachers," *Frontiers in Psychology*, vol. 14, p. 1272548, Sep. 2023, doi: 10.3389/fpsyg.2023.1272548.

[26] F. Habibzadeh, "Data distribution: normal or abnormal?" *Journal of Korean Medical Science*, vol. 39, no. 3, 2024, doi: 10.3346/jkms.2024.39.e35.

[27] W. R. B. Meredith D. Gall, Joyce P. Gall, *Educational research: an introduction*. 8th ed, Pearson/Allyn & Bacon, 2007, 2007.

[28] K. S. Taber, "The use of Cronbach's alpha when developing and reporting research instruments in science education," R*esearch in Science Education,* vol. 48, pp. 1273–1296, 2018, doi: 10.1007/s11165-016-9602-2.

[29] Á. González-Sanzana, J. Valenzuela-Carreño, P. Cáceres-Serrano, and L. Valdenegro-Fuentes, "Psychometric properties of the fit-choice scale factors influencing the choice of education careers in Chilean pedagogy students," *Electronic Journal of Research in Education Psychology*, vol. 22, no. 62, pp. 213–236, Apr. 2024, doi: 10.25115/ejrep.v22i62.9100.

[30] H. You, S. Park, M. Hong, and A. Warren, "Unveiling effectiveness: a meta-analysis of professional development programs in science education," *Journal of Research in Science Teaching*, vol. 62, no. 4, pp. 971–1005, 2025, doi: 10.1002/tea.21985.

[31] H. Jiang, R. Chugh, X. Zhai, K. Wang, and X. Wang, "Longitudinal analysis of teacher self-efficacy evolution during a STEAM professional development program: a qualitative case study," *Humanities and Social Sciences Communications*, vol. 11, p. 1162, 2024, doi: 10.1057/s41599-024-03655-5.
[32] S. Duan, K. Bissaker, and Z. Xu, "Correlates of teachers' classroom management self-efficacy: a systematic review and meta-analysis," *Educational Psychology Review*, vol. 36, no. 2, 2024, doi: 10.1007/s10648-024-09881-2.
[33] L. F. N. Jiménez, Y. A. O. Gomez, A. R. Maza, and J. E. F. Chiun, "The effectiveness of virtual learning environments in basic higher education: Analysis of their pedagogical impact," (in Spanish), *Sapiens in Education*, vol. 2, no. 2, pp. 1–20, Feb. 2025, doi: 10.71068/9183aj36.
[34] M. Ehlert and E. Souvignier, "Mind the gap: how teachers' professional development preferences relate to scientific recommendations," *Sage Open*, vol. 14, no. 3, 2024, doi: 10.1177/21582440241276172.
[35] P. Membiela, K. Acosta, and A. González, "Antecedents of Chilean in-service science teachers' emotions: the important role of teacher motivation towards teaching career and self-efficacy," *European Journal of Teacher Education*, vol. 48, no. 3, pp. 563–582, 2023, doi: 10.1080/02619768.2023.2229003.
[36] J. Liu, K. Wang, and Z. Pan, "The effectiveness of professional development in the self-efficacy of in-service teachers in STEM education: a meta-analysis," *Behavioral Sciences*, vol. 15, no. 10, p. 1364, Oct. 2025, doi: 10.3390/bs15101364.
[37] X. Zhou, L. Shu, Z. Xu, and Y. Padrón, "The effect of professional development on in-service STEM teachers' self-efficacy: a meta-analysis of experimental studies," *International Journal of STEM Education*, vol. 10, no. 1, May 2023, doi: 10.1186/s40594-023-00422-x.
[38] Y.-F. Yang, C. C. Tseng, and S.-C. Lai, "Enhancing teachers' self-efficacy beliefs in AI-based technology integration into English speaking teaching through a professional development program," *Teaching and Teacher Education*, vol. 144, p. 104582, 2024, doi: 10.1016/j.tate.2024.104582.
[39] B. Dreer, "On the outcomes of teacher wellbeing: a systematic review of research," *Frontiers in Psychology*, vol. 14, p. 1205179. 2023, doi: 10.3389/fpsyg.2023.1205179.
[40] M. Romero-Ariza, A. Quesada-Armenteros, and A. Estepa-Castro, "Promoting critical thinking through mathematics and science teacher education: The case of argumentation and graphs interpretation about climate change," *European Journal of Teacher Education*, vol. 47, no. 1, pp. 41–59, 2024, doi: 10.1080/02619768.2021.1961736.

BIOGRAPHIES OF AUTHORS

Katherine Acosta-García is an assistant professor in the Faculty of Education and Humanities at Universidad de Tarapacá, Chile. She received her doctoral degree in Education and Behavioral Sciences with a specialization in Innovation and Educational Research in Science Teaching from Universidad de Vigo, Spain (2021, Cum Laude), her master's degree in Didactics for Higher Education from Universidad de Tarapacá, Chile (2016), and her master's degree in Teaching Natural Sciences from Universidad Pedagógica Nacional, Colombia (2012). She holds a bachelor's degree in Biology from Universidad Pedagógica Nacional, Colombia (2010). In 2016, she joined the Department of Education at the Faculty of Education and Humanities of Universidad de Tarapacá, Arica, Chile, as an assistant professor. She has written several papers in the areas of science teaching, affective and motivational components in science teaching and learning, initial and continuing teacher training, and scientific inquiry. Her research interests also include teacher self-efficacy, emotions in science teaching, initial teacher training, scientific inquiry for science teaching, among others. She can be contacted at email: kacostag@academicos.uta.cl.

Marlene Morales Choque is a secondary school teacher in chemistry and holds a master's degree in Physical Chemistry. She serves as an academic at the University of Tarapacá, Chile, where she carries out teaching and research activities in the field of science education. Her academic trajectory has focused on science education, with particular interest in the analysis of pedagogical practices and the strengthening of teaching and learning processes in the sciences. She has developed work related to teacher education and reflective approaches to science teaching across different educational contexts, contributing to the improvement of educational quality in this area. Her research interests include science education, initial and continuing teacher education, and the development of teaching methodologies that promote meaningful learning in the sciences. She can be contacted at email: marlenemoralesicec@gmail.com.

Juan Pablo Aguirre Gonzalez is a mathematical engineer from the Universidad de Tarapacá and holds a master's degree in Statistics from the Universidad del Bío-Bío. He currently serves as a data analyst at Mesa Dinero SpA, where he specializes in the estimation of macroeconomic time series, focusing on Year-over-Year (YoY) and Quarter-over-Quarter (QoQ) growth rates for GDPC1, as well as providing monthly reports on the Consumer Price Index (CPIAUCNS). His research interests lie in the application of advanced statistical modeling to economic and social data. Regarding his professional achievements, he co-developed a "goodness-of-fit test for a bivariate negative binomial distribution" consistent against fixed alternatives, published in the journal Axioms (DOI: 10.3390/axioms14010054). He can be contacted at email: jpaguirregonzalez@gmail.com.

Sandra Ramírez Hernández is a psychopedagogue and doctoral candidate in Education at Universidad de Tarapacá, Chile. She received her specialist degree in Teaching in Psychopedagogy (2024), her master's degree in Special Education (2022), and her bachelor's degree in Education with a specialization in Pedagogy-Psychology (2016) from Universidad "Hermanos Saíz Montes de Oca", Pinar del Río, Cuba. From 2017 to 2024, she served as methodologist for preventive work at the Provincial Directorate of Education in Pinar del Río and as Adjunct Professor at Universidad "Hermanos Saíz Montes de Oca", teaching courses related to preventive work, hygiene of pedagogical activity, and psychopedagogical counseling. She has written several papers in the areas of educational inclusion, attention to students with special educational needs, and drug use prevention. Her research interests also include teacher self-efficacy, initial teacher training, comprehensive psychopedagogical attention, inclusive education, and educational prevention, among others. She can be contacted at email: sandrarh435@gmail.com.

Sara Maldonado-Araya is a nurse and a doctoral candidate in Education at the University of Tarapacá, Chile, and a doctoral candidate in the Doctoral Program in Communication at the University of Huelva, Spain. Her academic trajectory has developed within the field of higher education and educational research, with particular interest in university-level training processes in health-related academic programs. Her scholarly work focuses on the integration of digital technologies in higher education, particularly artificial intelligence, with emphasis on teacher beliefs, pedagogical practices, and institutional frameworks regulating technological adoption in university contexts. She has participated in research employing both quantitative and qualitative approaches, including bibliometric analyses, systematic reviews, and empirical studies in higher education. Her research interests include artificial intelligence in higher education, educational innovation, pedagogical practices, and university governance, especially in the context of health-related academic programs. She can be contacted at email: sarikaenf@gmail.com.