

## Article

# Performance on Standardized Pedagogical Assessments: A Sustainability-Oriented Analysis in Physical Education Teacher Education

Francisco Gallardo-Fuentes <sup>1,2,\*</sup> , Bastian Carter-Thuillier <sup>1,3</sup> , Johan Rivas-Valenzuela <sup>4</sup> ,  
Carolina Martínez-Angulo <sup>5</sup>, Sebastián Peña-Troncoso <sup>6,7</sup> , Jorge Gallardo-Fuentes <sup>8</sup> and Luis Añazco-Martínez <sup>9</sup> 

<sup>1</sup> Departamento de Educación, Universidad de los Lagos, Osorno 5311157, Chile; bastian.carter@ulagos.cl

<sup>2</sup> Facultad de Educación y Humanidades, Universidad Andrés Bello, Santiago 6513491, Chile

<sup>3</sup> Departamento de Didáctica y Práctica, Universidad Católica de Temuco, Temuco 4810399, Chile

<sup>4</sup> Group EFyM, Physical Education School, Pontificia Universidad Católica de Valparaíso, Valparaíso 2340000, Chile; johan.rivas@pucv.cl

<sup>5</sup> Departamento de Ciencias de la Actividad Física, Universidad de los Lagos, Osorno 5311157, Chile; carolina.martinez@ulagos.cl

<sup>6</sup> Facultad Educación, Universidad Austral de Chile, Valdivia 5110693, Chile; sebastian.pena@uach.cl

<sup>7</sup> Facultad de Educación y Cultura, Universidad SEK, Santiago 7520317, Chile

<sup>8</sup> Escuela de Medicina Veterinaria, Facultad de Medicina y Ciencias de la Salud, Universidad Mayor, Temuco 4791297, Chile; jorgegallardofuentes@gmail.com

<sup>9</sup> Facultad de Ciencias Humanas, Universidad Arturo Prat, Iquique 1110939, Chile; lanazco@unap.cl

\* Correspondence: fgallardo@ulagos.cl or f.gallardofuentes@uandresbello.eu; Tel.: +56-9642333026

## Abstract

Initial teacher education requires assessment mechanisms capable of identifying strengths and weaknesses in professional preparation. This study analyzes the performance of students in Initial Teacher Education in Physical Education against the pedagogical standards assessed by the 2024 National Diagnostic Assessment (NDA) in Chile, examining achievement patterns and differences by sex and administrative dependency. A quantitative, cross-sectional, and correlational design was employed, based on official census data corresponding to 1035 students enrolled in Chilean universities offering this degree program. Descriptive analyses and non-parametric inferential tests were conducted, using the Mann–Whitney U test for comparisons by sex, and the Kruskal–Wallis test followed by Dunn’s post hoc comparisons with Bonferroni correction for comparisons across administrative dependency categories. The highest levels were observed in professional ethics ( $M = 0.69$ ) and student learning and development ( $M = 0.57$ ), and the lowest in respectful and well organized learning environments ( $M = 0.39$ ) and curricular pedagogical content knowledge ( $M = 0.40$ ). Significant differences by sex appeared only in personal and social development and professional ethics ( $p = 0.01$ ;  $r = 0.08$ ), favoring women. Differences by administrative dependency were significant in few standards and consistently small ( $r \leq 0.14$ ). These effects indicate that achievement is largely comparable across institutions and that the main weaknesses are systemic. Strengthening these dimensions constitutes a challenge for the sustainability of teacher education, rather than a problem restricted to specific institutions.

**Keywords:** initial teacher education; physical education; diagnostic assessment; educational sustainability



Academic Editor: Antonio P. Gutierrez de Blume

Received: 19 June 2026

Revised: 20 July 2026

Accepted: 21 July 2026

Published: 30 July 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 \(CC BY\)](#) license.

## 1. Introduction

In the context of contemporary educational transformations, initial teacher education (ITE) has been recognized as a strategic component for advancing toward sustainable, inclusive, and high-quality education systems. For this reason, institutions such as the World Bank allocate funding to support this educational stage in developing countries [1]. In particular, the framework of Education for Sustainable Development (ESD), promoted by UNESCO, has emphasized the need to prepare teachers capable of addressing complex and multidimensional challenges, such as inequality, climate change, and social transformation, in alignment with the Sustainable Development Goals (especially SDG 4: Quality Education) [2]. From this perspective, sustainable teacher education implies not only the development of pedagogical competencies, but also the capacity to promote meaningful, transferable, and sustainable learning over time (long-term learning outcomes) [3].

In this study, educational sustainability is understood as the capacity of teacher education systems to generate durable, transferable, and equitable professional learning over time, in line with the ESD framework and SDG 4 [2]. From this perspective, pedagogical standards and national diagnostic assessments operate as monitoring instruments for the sustainability of teacher education, providing system-wide evidence on whether programs consolidate the competencies that enable graduates to promote lasting learning, inclusion, and well-being [3,4]. This link between pedagogical benchmarks and long-term educational sustainability frames the analysis presented here.

Within this scenario, the implementation of ITE assessment systems plays a fundamental role as a quality assurance mechanism and as a tool for strengthening systemic coherence between educational policies, training processes, and learning outcomes [4]. The literature highlights that sustainable education systems require not only clear standards, but also assessment mechanisms that make it possible to monitor the development of teaching competencies in a valid and relevant manner [5].

However, the use of these assessments raises important tensions from the perspective of educational sustainability. Several studies warn that although these instruments make it possible to measure outcomes on a large scale, they tend to privilege more easily quantifiable dimensions, which may limit the understanding of complex, contextualized, and situated competencies that are essential for teaching practice [6,7]. Consequently, authors have pointed out the need to move toward more comprehensive assessment approaches capable of capturing not only immediate performance, but also the ability of pre-service teachers to foster sustainable learning and adapt to changing contexts [8].

In the Chilean case, the National Diagnostic Assessment (NDA) of Initial Teacher Education is configured as a key instrument within the quality assurance system for teacher education [9]. Census-based and mandatory in nature, this assessment makes it possible to measure the level of achievement of the teaching profession standards defined by the Ministry of Education, distinguishing between disciplinary and pedagogical standards. From a sustainability perspective, the NDA offers a valuable opportunity to analyze the coherence between training processes and observed outcomes, thereby contributing to an understanding of the strengths and weaknesses of the teacher education system [10]. However, it also raises questions regarding its ability to reflect the complexity of teacher learning and its potential contribution to the development of sustainable competencies [11].

This analysis is particularly relevant in the field of Physical Education (PE), a discipline that presents distinctive characteristics in teacher education [12]. PE lies at the intersection of bodily, cognitive, social, and emotional dimensions, making it a privileged space for promoting well-being, inclusion, and social sustainability [13]. In line with the principles of Education for Sustainable Development (ESD), this discipline has the potential to contribute to the development of active and healthy lifestyles, as well as to the formation of

critical citizens committed to their environment [14]. However, this complexity also poses significant challenges for its assessment, particularly with regard to measuring pedagogical competencies that are developed in highly situated contexts [15].

Within this framework, pedagogical standards play a central role as a reference point for guiding teacher education toward sustainable practices. These standards include competencies related to planning, teaching implementation, understanding students, and the assessment of learning, all of which are fundamental to the development of effective and contextualized teaching [16]. From the perspective of educational sustainability, these competencies should be understood as comprehensive capacities that enable teachers to act in a reflective, ethical, and adaptive manner, thereby contributing to the generation of lasting and meaningful learning for students [17].

However, recent evidence suggests the existence of a persistent gap between training processes in initial teacher education (ITE) and the results observed in standardized assessments, particularly in the area of pedagogical competencies [18]. This gap poses significant challenges to the sustainability of the education system, as it calls into question the capacity of teacher education to generate learning that endures over time and translates into effective pedagogical practices. In this regard, it has been pointed out that the lack of coherence among teacher education, assessment, and teaching practice may limit the impact of educational policies and hinder the consolidation of sustainable teacher education systems [19].

Despite the relevance of these challenges, a gap remains in the literature regarding the disaggregated analysis of performance on pedagogical standards in standardized teacher education assessments [20]. Most studies focus on overall results, the analysis of disciplinary standards, or broad comparisons, making it difficult to identify specific performance patterns and to understand differences across distinct components of pedagogical preparation [21,22]. This gap is particularly evident in Latin American contexts, where assessment systems have undergone recent transformations within the framework of quality assurance policies [23].

Two additional dimensions warrant disaggregated attention. First, sex-based analyses are pertinent because international evidence has consistently documented gender differences in educational achievement and in socioemotional dispositions, and because PE remains a field historically shaped by gendered practices and expectations; examining whether pedagogical standards are attained equally by women and men is therefore an equity question directly aligned with SDG 4 and SDG 5 [2]. Second, administrative dependency (public universities, traditional private universities belonging to the G9 group, and private universities) constitutes a structural feature of Chile's diversified higher education system; comparing achievement across these categories makes it possible to assess whether teacher education outcomes are equitably distributed across institutional types—an issue central to the social sustainability of the system [9,19].

In line with the above, the aim of this article is to analyze the performance of students in Physical Education Teacher Education (PETE) across the different pedagogical standards assessed in the NDA, using census data corresponding to the year 2024, and to examine differences in achievement according to sex and to the administrative dependency of the institutions.

## 2. Materials and Methods

### 2.1. Study Design and Participants

The study was conducted using a quantitative approach, with a cross-sectional design and correlational scope, based on the secondary analysis of data derived from the 2024 NDA [24]. The dataset corresponds to the complete national census of students who

took the 2024 NDA in PE teacher education program in Chile ( $N = 1035$ ; 304 women and 731 men), enrolled in the 29 universities that offered this degree program at the time of the assessment. No institutions offering the specialization were excluded and all students with valid records in the official database of the Ministry of education were included in the analysis; accordingly, the data represent the entire national universe of PETE students assessed in 2024 rather than a sample.

## 2.2. Instrument and Procedure: The National Diagnostic Assessment

The NDA is a mandatory, census-based assessment established by Law No. 20.903 and administered annually by the Center for Pedagogical Improvement, Experimentation and Research (CPEIP) of the Chilean Ministry of Education. All students enrolled in accredited teacher education programs must take it during their penultimate year of study, and its results are reported to the universities to inform curricular improvement processes [9]. The instrument includes a standardized test of pedagogical knowledge composed of closed, multiple-choice items referenced to the pedagogical standards for teaching careers; items are developed, piloted, and psychometrically validated by CPEIP for application in the Chilean context. Achievement is operationalized as the percentage of correct responses obtained by each student in the set of items associated with each pedagogical standard, expressed as a proportion ranging from 0 to 1. Consequently, the means and standard deviations reported in the tables of this article correspond to averages of these proportions of achievement; no course grades, student surveys, or external institutional evaluations were involved in the measurement.

Specifically, the analysis focused on the pedagogical standards that serve as national benchmarks for pedagogical performance, outlining the knowledge, skills, and dispositions that prospective teachers are expected to demonstrate throughout their initial teacher education [25]. The current national framework for teaching careers, approved in 2021, comprises twelve pedagogical standards organized into four domains. However, the official 2024 NDA results database reports achievement data for nine of them, given that Standard 9 (Assessment and Feedback for Learning), Standard 11 (Continuous Professional Learning), and Standard 12 (Commitment to the Continuous Improvement of the School Community) refer to competencies enacted in situated professional practice, which are not measured through the written diagnostic instrument. The nine standards analyzed are distributed as follows: (a) Standard 1: Student Learning and Development; (b) Standard 2: Pedagogical Content Knowledge and Knowledge of the School Curriculum; (c) Standard 3: Teaching Planning; (d) Standard 4: Assessment Planning; (e) Standard 5: Respectful and Organized Learning Environment; (f) Standard 6: Personal and Social Development; (g) Standard 7: Teaching Strategies for the Achievement of Deep Learning; (h) Standard 8: Learning Strategies for the Development of Thinking Skills; (i) Standard 10: Professional Ethics.

## 2.3. Data Analysis

Data processing was carried out using R software 4.6.1 [26], with the packages dplyr, tidyr, readxl, and openxlsx for database management, cleaning, analysis, and export. Initially, a descriptive analysis was performed, considering arithmetic means ( $M$ ), standard deviations ( $SD$ ), and frequencies, according to the characteristics of the analyzed data [26], followed by inferential analysis. Given the nature of the data proportions of achievement bounded between 0 and 1, which cannot be assumed to behave as normally distributed interval measures non-parametric procedures were considered the appropriate analytical choice; this decision was empirically corroborated by the Shapiro–Wilk test, which indicated significant deviations from normality in all variables. Accordingly, the Mann–Whitney U test was used to compare the two sex groups (women vs. men), whereas comparisons across the

three administrative dependency categories (public, private G9, and private universities) were analyzed through pairwise comparisons using the Mann–Whitney U test among the three institutional categories, considering all possible pairwise comparisons. To control the inflation of Type I error associated with multiple testing, Bonferroni correction was applied to the  $p$ -values. For each comparison, the U statistic, Bonferroni-adjusted  $p$ -value, and effect size ( $r$ ), estimated from the Z statistic, were reported. Statistical significance was set at  $p < 0.05$  for all analyses. For each comparison, the U statistic,  $p$ -value, and effect size  $r$ , estimated from the Z statistic, were reported to complement statistical significance. Following Cohen’s benchmarks,  $r$  values of approximately 0.10, 0.30, and 0.50 were interpreted as small, medium, and large effects, respectively [27]. The latter was assessed using a significance level of  $p < 0.05$ .

### 3. Results

Table 1 presents the descriptive results related to the percentage of achievement in each standard among the 1035 students enrolled in the different universities offering this teacher education specialization.

**Table 1.** Means and standard deviations of achievement in the pedagogical standards by university.

| University                                                | N    | Stan. 1 |      | Stan. 2 |      | Stan. 3 |      | Stan. 4 |      | Stan. 5 |      | Stan. 6 |      | Stan. 7 |      | Stan. 8 |      | Stan. 10 |      |
|-----------------------------------------------------------|------|---------|------|---------|------|---------|------|---------|------|---------|------|---------|------|---------|------|---------|------|----------|------|
|                                                           |      | M       | sd   | M       | sd   | M       | sd   | M       | sd   | M       | sd   | M       | sd   | M       | sd   | M       | sd   | M        | sd   |
| The Pontifical Catholic University of Chile               | 24   | 0.72    | 0.18 | 0.55    | 0.26 | 0.60    | 0.18 | 0.54    | 0.22 | 0.60    | 0.31 | 0.59    | 0.18 | 0.57    | 0.22 | 0.64    | 0.23 | 0.77     | 0.20 |
| The Pontifical Catholic of Valparaiso University of Chile | 7    | 0.61    | 0.21 | 0.50    | 0.29 | 0.52    | 0.18 | 0.64    | 0.28 | 0.19    | 0.18 | 0.44    | 0.31 | 0.47    | 0.32 | 0.53    | 0.21 | 0.61     | 0.21 |
| University Adventist of Chile                             | 7    | 0.50    | 0.20 | 0.21    | 0.30 | 0.55    | 0.19 | 0.57    | 0.19 | 0.24    | 0.16 | 0.52    | 0.19 | 0.55    | 0.05 | 0.57    | 0.22 | 0.69     | 0.28 |
| University Andres Bello                                   | 58   | 0.62    | 0.19 | 0.44    | 0.21 | 0.51    | 0.25 | 0.59    | 0.23 | 0.32    | 0.27 | 0.53    | 0.18 | 0.48    | 0.22 | 0.51    | 0.20 | 0.72     | 0.20 |
| University Arturo Prat                                    | 17   | 0.64    | 0.21 | 0.25    | 0.22 | 0.53    | 0.19 | 0.59    | 0.28 | 0.37    | 0.23 | 0.51    | 0.15 | 0.50    | 0.16 | 0.54    | 0.21 | 0.61     | 0.19 |
| University Austral of Chile                               | 46   | 0.55    | 0.17 | 0.40    | 0.20 | 0.43    | 0.19 | 0.51    | 0.25 | 0.27    | 0.22 | 0.48    | 0.17 | 0.47    | 0.18 | 0.50    | 0.20 | 0.66     | 0.20 |
| University Autonoma de Chile                              | 101  | 0.53    | 0.19 | 0.35    | 0.20 | 0.48    | 0.21 | 0.49    | 0.24 | 0.42    | 0.28 | 0.47    | 0.18 | 0.43    | 0.21 | 0.49    | 0.20 | 0.63     | 0.22 |
| University Bernardo O’higgins                             | 14   | 0.58    | 0.21 | 0.39    | 0.31 | 0.54    | 0.23 | 0.54    | 0.24 | 0.36    | 0.21 | 0.45    | 0.14 | 0.47    | 0.22 | 0.58    | 0.21 | 0.78     | 0.17 |
| University Catholic Silva Henriquez                       | 52   | 0.55    | 0.18 | 0.42    | 0.25 | 0.51    | 0.21 | 0.58    | 0.23 | 0.42    | 0.26 | 0.52    | 0.14 | 0.39    | 0.20 | 0.51    | 0.18 | 0.66     | 0.21 |
| University Catholic of Temuco                             | 27   | 0.63    | 0.15 | 0.28    | 0.21 | 0.59    | 0.20 | 0.67    | 0.27 | 0.43    | 0.29 | 0.48    | 0.13 | 0.46    | 0.19 | 0.61    | 0.20 | 0.67     | 0.21 |
| University Catholic of Santisima Concepcion               | 71   | 0.57    | 0.19 | 0.32    | 0.23 | 0.54    | 0.19 | 0.54    | 0.21 | 0.39    | 0.32 | 0.41    | 0.18 | 0.49    | 0.22 | 0.50    | 0.21 | 0.69     | 0.21 |
| University Catholic of Maule                              | 43   | 0.60    | 0.18 | 0.37    | 0.23 | 0.59    | 0.20 | 0.57    | 0.28 | 0.43    | 0.27 | 0.48    | 0.16 | 0.49    | 0.20 | 0.51    | 0.18 | 0.66     | 0.18 |
| University Central de Chile                               | 22   | 0.61    | 0.15 | 0.38    | 0.29 | 0.42    | 0.17 | 0.53    | 0.25 | 0.47    | 0.28 | 0.49    | 0.11 | 0.46    | 0.20 | 0.59    | 0.20 | 0.68     | 0.18 |
| University Mayor                                          | 49   | 0.61    | 0.17 | 0.43    | 0.23 | 0.59    | 0.19 | 0.65    | 0.24 | 0.37    | 0.28 | 0.58    | 0.17 | 0.52    | 0.19 | 0.58    | 0.18 | 0.76     | 0.16 |
| University Metropolitan Educational Sciences              | 44   | 0.61    | 0.18 | 0.42    | 0.21 | 0.57    | 0.21 | 0.57    | 0.24 | 0.52    | 0.32 | 0.49    | 0.13 | 0.52    | 0.22 | 0.54    | 0.20 | 0.74     | 0.20 |
| University San Sebastian                                  | 34   | 0.51    | 0.19 | 0.46    | 0.27 | 0.50    | 0.23 | 0.58    | 0.22 | 0.38    | 0.26 | 0.45    | 0.16 | 0.46    | 0.23 | 0.51    | 0.19 | 0.73     | 0.22 |
| University Santo Tomas                                    | 85   | 0.58    | 0.17 | 0.43    | 0.25 | 0.61    | 0.21 | 0.59    | 0.25 | 0.38    | 0.28 | 0.53    | 0.18 | 0.45    | 0.19 | 0.56    | 0.20 | 0.73     | 0.18 |
| University of Antofagasta                                 | 29   | 0.59    | 0.22 | 0.39    | 0.24 | 0.48    | 0.20 | 0.65    | 0.25 | 0.44    | 0.27 | 0.45    | 0.15 | 0.41    | 0.21 | 0.55    | 0.22 | 0.69     | 0.21 |
| University of Atacama                                     | 14   | 0.54    | 0.24 | 0.54    | 0.31 | 0.54    | 0.20 | 0.41    | 0.25 | 0.36    | 0.24 | 0.58    | 0.20 | 0.45    | 0.22 | 0.54    | 0.16 | 0.60     | 0.23 |
| University of Concepcion                                  | 34   | 0.64    | 0.16 | 0.35    | 0.25 | 0.53    | 0.19 | 0.52    | 0.21 | 0.43    | 0.32 | 0.49    | 0.15 | 0.42    | 0.15 | 0.47    | 0.19 | 0.73     | 0.17 |
| University la Frontera                                    | 32   | 0.52    | 0.21 | 0.50    | 0.26 | 0.55    | 0.17 | 0.54    | 0.25 | 0.48    | 0.35 | 0.46    | 0.18 | 0.52    | 0.21 | 0.48    | 0.24 | 0.74     | 0.15 |
| University of las Americas                                | 12   | 0.66    | 0.16 | 0.42    | 0.27 | 0.46    | 0.18 | 0.71    | 0.14 | 0.31    | 0.17 | 0.44    | 0.15 | 0.42    | 0.22 | 0.48    | 0.18 | 0.62     | 0.16 |
| University of los Lagos                                   | 57   | 0.50    | 0.20 | 0.42    | 0.27 | 0.46    | 0.22 | 0.52    | 0.22 | 0.30    | 0.26 | 0.41    | 0.15 | 0.37    | 0.18 | 0.48    | 0.20 | 0.61     | 0.22 |
| University of Magallanes                                  | 4    | 0.50    | 0.27 | 0.31    | 0.13 | 0.33    | 0.14 | 0.31    | 0.31 | 0.58    | 0.50 | 0.36    | 0.27 | 0.43    | 0.26 | 0.46    | 0.18 | 0.64     | 0.08 |
| University of Playa Ancha                                 | 45   | 0.55    | 0.16 | 0.42    | 0.23 | 0.54    | 0.19 | 0.50    | 0.24 | 0.40    | 0.26 | 0.52    | 0.17 | 0.46    | 0.17 | 0.44    | 0.23 | 0.71     | 0.17 |
| University of Santiago de Chile                           | 26   | 0.61    | 0.20 | 0.43    | 0.21 | 0.56    | 0.17 | 0.56    | 0.22 | 0.46    | 0.25 | 0.54    | 0.14 | 0.46    | 0.24 | 0.52    | 0.21 | 0.69     | 0.20 |
| University of Tarapaca                                    | 35   | 0.52    | 0.20 | 0.29    | 0.20 | 0.45    | 0.17 | 0.47    | 0.24 | 0.30    | 0.25 | 0.42    | 0.15 | 0.33    | 0.20 | 0.38    | 0.24 | 0.62     | 0.19 |
| University of Viña del Mar                                | 6    | 0.58    | 0.13 | 0.33    | 0.20 | 0.53    | 0.22 | 0.50    | 0.27 | 0.22    | 0.27 | 0.45    | 0.18 | 0.40    | 0.28 | 0.45    | 0.28 | 0.64     | 0.15 |
| University of Bio Bio                                     | 40   | 0.58    | 0.15 | 0.39    | 0.27 | 0.48    | 0.18 | 0.53    | 0.23 | 0.32    | 0.26 | 0.49    | 0.15 | 0.40    | 0.21 | 0.48    | 0.19 | 0.66     | 0.21 |
| National total                                            | 1035 | 0.57    | 0.19 | 0.40    | 0.25 | 0.53    | 0.21 | 0.56    | 0.24 | 0.39    | 0.28 | 0.49    | 0.17 | 0.46    | 0.21 | 0.51    | 0.21 | 0.69     | 0.20 |

Stan. = Standard; M. = Means; sd. = Standard Deviation.

Considering the overall average across standards, Pontifical University Católica de Chile shows the highest levels of achievement, with mean scores consistently above those of the other institutions and standing out particularly in Standards 1, 7, 8, and 10. Favorable results are also observed at University Mayor, Universidad Metropolitana Educational Sciences, and University Santo Tomás, whose overall averages rank among the highest in the sample. In contrast, the lowest average scores correspond to University of Tarapacá, University of Magallanes, and University of Los Lagos, reflecting relatively lower performance across several of the standards analyzed. When examining the standards individually, the highest levels of achievement tend to be concentrated in Standards 7, 8, and 10, where

several universities reach mean scores above 0.60. Pontifical University Católica of Chile and University of Chile stand out in particular, as they present some of the highest values in the entire table.

By contrast, Standards 4 and 5 generally show the lowest mean scores, thus emerging as the areas posing the greatest difficulties for the participating institutions. This pattern suggests an uneven distribution of performance across standards, with some areas showing more consolidated achievement and others representing common challenges for a large proportion of the universities. Regarding the dispersion of the results, the standard deviations show moderate values in most institutions and standards, indicating relatively low internal variability. Table 2 presents the percentages of achievement by standard for men and women.

**Table 2.** Means and standard deviations of achievement by standard, women vs. men.

| Pedagogical Standard                                                    | Women (N = 304) |      | Men (n = 731) |      | W vs. M   |        |      |
|-------------------------------------------------------------------------|-----------------|------|---------------|------|-----------|--------|------|
|                                                                         | M               | sd   | M             | sd   | U         | p      | r    |
| 1: Student Learning and Development                                     | 0.58            | 0.19 | 0.57          | 0.19 | 112,718   | 0.71   | 0.01 |
| 2: Pedagogical Content Knowledge and Knowledge of the School Curriculum | 0.40            | 0.26 | 0.39          | 0.23 | 113,106   | 0.63   | 0.01 |
| 3: Teaching Planning                                                    | 0.54            | 0.22 | 0.52          | 0.20 | 116,262.5 | 0.23   | 0.04 |
| 4: Assessment Planning                                                  | 0.57            | 0.25 | 0.55          | 0.24 | 116,369   | 0.21   | 0.04 |
| 5: Respectful and Organized Learning Environment                        | 0.41            | 0.29 | 0.38          | 0.28 | 114,770   | 0.37   | 0.03 |
| 6: Personal and Social Development                                      | 0.51            | 0.16 | 0.48          | 0.17 | 122,524   | 0.01 * | 0.08 |
| 7: Teaching Strategies for the Achievement of Deep Learning             | 0.46            | 0.22 | 0.45          | 0.20 | 114,650   | 0.41   | 0.03 |
| 8: Learning Strategies for the Development of Thinking Skill            | 0.52            | 0.22 | 0.51          | 0.20 | 113,850.5 | 0.52   | 0.02 |
| 9: Professional Ethics                                                  | 0.71            | 0.20 | 0.68          | 0.20 | 122,148   | 0.01 * | 0.08 |

M. = Means; sd. = Standard Deviation; U = Mann–Whitney U test; r = Effect Size; W vs. M = Comparison between Women and Men; \*  $p \leq 0.05$ .

Table 2 shows that women have slightly higher mean scores than men across all the standards assessed. The highest scores are observed in Standard 10 for both groups, whereas the lowest values are concentrated in Standards 2 and 5. In general, the standard deviations are similar for women and men, indicating a comparable dispersion of results. In inferential terms, statistically significant differences are observed only in Standard 6 ( $p = 0.01$ ;  $r = 0.08$ ) and Standard 10 ( $p = 0.01$ ;  $r = 0.08$ ), both in favor of women. However, the effect sizes are small, so the practical magnitude of these differences is limited. No significant differences between men and women were identified in the remaining standards.

Table 3 presents the results for each standard according to the administrative dependency of the university institutions. These are divided into three categories:

1. Universities with public administrative dependency: (University Arturo Prat; University of Antofagasta; University of Atacama; University la Frontera; University of los Lagos; University of Magallanes; University of Playa Ancha; University of Santiago de Chile; University of Tarapaca; University of Bio Bio; University Metropolitan Educational Sciences).
2. Traditional private universities belonging to the G9 (Pontifical University Católica of Chile; Pontifical University Católica of Valparaíso; University Austral of Chile; University Catholic of Temuco; University Catholic of Maule; University of Concepción),
3. Private universities: (University Adventista de Chile; University Andres Bello; University Autónoma de Chile; University Bernardo O’Higgins; University Catholic of Santísima Concepción; University Catholic Silva Henríquez; University Central de Chile; University of las Américas; University of Viña del Mar; University Mayor; University San Sebastián; University Santo Tomás).

Table 3 shows heterogeneous performance for the total sample. Overall, the mean scores range from 0.39 to 0.69, revealing relevant descriptive differences in the level of achievement attained depending on the standard analyzed. This variation supports the view that pedagogical performance includes areas of relatively greater consolidation and

others with lower results. The highest mean score is observed in Standard 10 ( $M = 0.67$ ;  $sd = 0.20$ ), followed by Standard 1 ( $M = 0.57$ ;  $sd = 0.19$ ) and Standard 4 ( $M = 0.55$ ;  $sd = 0.24$ ). These results suggest that students show relatively better performance in these pedagogical components. At an intermediate level are Standard 3 ( $M = 0.52$ ;  $sd = 0.21$ ), Standard 8 ( $M = 0.51$ ;  $sd = 0.21$ ), Standard 6 ( $M = 0.49$ ;  $sd = 0.17$ ), and Standard 7 ( $M = 0.45$ ;  $sd = 0.21$ ), whose results indicate moderate levels of achievement within the overall set. The lowest mean scores are concentrated in Standard 5 ( $M = 0.39$ ;  $sd = 0.28$ ) and Standard 2 ( $M = 0.40$ ;  $sd = 0.24$ ), which appear as the areas of greatest weakness in teacher preparation.

**Table 3.** Means and Standard deviations of achievement according to administrative dependency.

| Pedagogical Standard                                                    | Sample Total<br>(N = 1035) |      | PR<br>(N = 511) |      | PG9<br>(N = 181) |      | PU<br>(N = 343) |      |
|-------------------------------------------------------------------------|----------------------------|------|-----------------|------|------------------|------|-----------------|------|
|                                                                         | M                          | sd   | M               | sd   | M                | sd   | M               | sd   |
| 1: Student Learning and Development                                     | 0.57                       | 0.19 | 0.57            | 0.18 | 0.62             | 0.18 | 0.56            | 0.20 |
| 2: Pedagogical Content Knowledge and Knowledge of the School Curriculum | 0.40                       | 0.24 | 0.39            | 0.24 | 0.39             | 0.24 | 0.41            | 0.25 |
| 3: Teaching Planning                                                    | 0.52                       | 0.21 | 0.53            | 0.21 | 0.54             | 0.20 | 0.51            | 0.20 |
| 4: Assessment Planning                                                  | 0.55                       | 0.24 | 0.57            | 0.24 | 0.56             | 0.25 | 0.53            | 0.24 |
| 5: Respectful and Organized Learning Environment                        | 0.39                       | 0.28 | 0.39            | 0.28 | 0.40             | 0.29 | 0.39            | 0.29 |
| 6: Personal and Social Development                                      | 0.49                       | 0.17 | 0.49            | 0.17 | 0.50             | 0.17 | 0.48            | 0.16 |
| 7: Teaching Strategies for the Achievement of Deep Learning             | 0.45                       | 0.21 | 0.46            | 0.21 | 0.48             | 0.20 | 0.43            | 0.21 |
| 8: Learning Strategies for the Development of Thinking Skill            | 0.51                       | 0.21 | 0.53            | 0.20 | 0.53             | 0.21 | 0.49            | 0.21 |
| 9: Professional Ethics                                                  | 0.69                       | 0.20 | 0.69            | 0.20 | 0.69             | 0.20 | 0.67            | 0.20 |

M. = Means; sd. = Standard Deviation; PR = Private University; PG9 = Private University G9; PU = Public University.

Across administrative categories, several standards are very close to one another, as is the case for Standard 2, where the three categories range between 0.39 and 0.41; Standard 5, where the values fall between 0.39 and 0.40; and Standard 10, where the mean scores range from 0.67 to 0.69. Although small variations exist between groups, performance by administrative dependency tends to maintain a relatively similar profile, with more visible differences in some specific standards, especially Standard 1 and Standard 8. G9 private institutions show the highest mean scores in several standards, particularly in Standard 1 ( $M = 0.62$ ), Standard 3 ( $M = 0.54$ ), Standard 5 ( $M = 0.40$ ), Standard 6 ( $M = 0.50$ ), Standard 7 ( $M = 0.48$ ), and Standard 8 ( $M = 0.53$ ). For their part, private institutions show the highest mean scores in Standard 4 ( $M = 0.57$ ) and Standard 10 ( $M = 0.69$ ), whereas public institutions reach the highest relative mean only in Standard 2 ( $M = 0.41$ ). This pattern suggests a slight descriptive advantage for G9 private institutions across a substantial proportion of the pedagogical standards analyzed. From an inferential perspective, comparisons between the percentages of achievement according to administrative dependency are presented in Table 4.

**Table 4.** Mean comparisons results by administrative dependency.

| Standard vs. Administrative Dependency                                  | PG9 vs. PR |        |      | PG9 vs. PU |        |      | PR vs. PU |        |      |
|-------------------------------------------------------------------------|------------|--------|------|------------|--------|------|-----------|--------|------|
|                                                                         | U          | Sig    | r    | U          | Sig    | r    | U         | Sig    | r    |
| 1: Student Learning and Development                                     | 52,764.5   | 0.01 * | 0.11 | 36,086.5   | 0.01 * | 0.14 | 89,914.5  | 0.51   | 0.02 |
| 2: Pedagogical Content Knowledge and Knowledge of the School Curriculum | 45,526.5   | 0.74   | 0.01 | 29,796     | 0.43   | 0.03 | 85,463.5  | 0.52   | 0.02 |
| 3: Teaching Planning                                                    | 47,915     | 0.46   | 0.03 | 33,574     | 0.11   | 0.07 | 91,403.5  | 0.27   | 0.04 |
| 4: Assessment Planning                                                  | 45,538.5   | 0.75   | 0.01 | 32,926.5   | 0.23   | 0.05 | 94,456.5  | 0.04 * | 0.07 |
| 5: Respectful and Organized Learning Environment                        | 47,890     | 0.45   | 0.03 | 31,679.5   | 0.68   | 0.02 | 86,332.5  | 0.69   | 0.01 |
| 6: Personal and Social Development                                      | 47,143.5   | 0.69   | 0.01 | 33,688     | 0.10   | 0.07 | 93,082    | 0.12   | 0.05 |
| 7: Teaching Strategies for the Achievement of Deep Learning             | 47,819     | 0.49   | 0.03 | 34,129     | 0.06   | 0.08 | 93,031.5  | 0.12   | 0.05 |
| 8: Learning Strategies for the Development of Thinking Skill            | 47,284.5   | 0.65   | 0.02 | 34,469.5   | 0.03 * | 0.09 | 95,544    | 0.02 * | 0.08 |
| 9: Professional Ethics                                                  | 44,785     | 0.52   | 0.02 | 31,967.5   | 0.57   | 0.03 | 93,005.5  | 0.12   | 0.05 |

PR = Private University; PG9 = Private University G9; PU = Public University; U = Mann–Whitney U test; r = Effect Size; \*  $p \leq 0.05$ .

Few of the descriptive differences observed reach statistical significance. In most comparisons across administrative dependency categories, the results indicate an absence

of statistically significant differences between the groups compared. This is especially evident in Standards 2, 3, 5, 6, 7, and 10, where none of the comparisons among G9 private, private, and public institutions reach statistical significance. Therefore, in these cases, the small variations in the mean scores should be interpreted as descriptive differences, but not as inequality across administrative categories. Regarding effect size, the  $r$  values are low in all comparisons, ranging from 0.01 to 0.14. This indicates that, even in those cases where statistically significant differences exist, the magnitude of such differences is small. Consequently, the results should be interpreted with caution, avoiding an overestimation of the practical relevance of the differences across administrative categories. Overall, the findings suggest specific, low-magnitude variations in certain pedagogical standards.

More specifically, statistically significant differences are concentrated in three standards. In Standard 1, differences are observed between G9 private and private institutions ( $p = 0.01$ ;  $r = 0.11$ ), and between G9 private and public institutions ( $p = 0.01$ ;  $r = 0.14$ ), with higher mean scores for G9 private institutions. In Standard 4, a significant difference appears between private and public institutions ( $p = 0.04$ ;  $r = 0.07$ ), with a higher mean score in private institutions the significant difference appears between private and public institutions ( $p = 0.04$ ;  $r = 0.07$ ), with a higher mean score in private institutions. Finally, in Standard 8, differences are identified between G9 private and public institutions ( $p = 0.03$ ;  $r = 0.09$ ), and between private and public institutions ( $p = 0.02$ ;  $r = 0.08$ ), again with lower mean scores in public institutions. Taken together, the statistically significant results are concentrated in comparisons in which public institutions tend to show lower mean scores, although always with small effect sizes.

#### 4. Discussion

The findings support the conclusion that performance, expressed as the percentage of achievement in the pedagogical standards of the NDA, does not reflect homogeneous achievement in Chilean Initial Teacher Education in PE, but rather a configuration of strengths and weaknesses that reveals different levels of consolidation of learning within this teacher education specialization. From the perspective of sustainable teacher education, and considering the heterogeneity of the results, this is particularly significant, since educational sustainability depends not only on the existence of formal standards, but also on future teachers developing consistent capacities to sustain inclusive, equitable, and pedagogically robust learning trajectories over time [2,16]. In this regard, initial teacher education has been described as a strategic component of the adaptive capacity of the school system to prepare citizens capable of engaging positively in society, precisely because it articulates normative expectations, pedagogical knowledge, and the projection of educational quality [1,4]. Therefore, the pattern observed in this study suggests that the preparation provided by Initial Teacher Education in PE in Chile shows relevant progress in some areas, but still exhibits training-related bottlenecks in others, especially where professional competence requires moving from conceptual understanding to situated pedagogical action, as has been noted both in the literature on education for sustainable development and in recent research on action competence in pre-service teachers [7,28].

Within this framework, the fact that relatively higher performance is concentrated in professional ethics and in standards associated with student learning and development may be interpreted as a sign that teacher education has succeeded in establishing a relatively solid normative repertoire regarding the public purpose of teaching. This interpretation is consistent with debates on the NDA in Chile, which have shown that the instrument tends to make more visible those dimensions that are more readily recognizable within the institutional discourse of teacher education [10,11]. However, interpreting these better performances as sufficient evidence of quality would be reductionist, since pedagogy is a

science that unfolds in practice. As Biesta [7], warns, teaching cannot be judged solely by adherence to ends that are recognized as valuable, but also by the way in which those ends are translated into pedagogical decisions. In PE, this distinction is crucial, as professional ethics only acquires formative rigor when it is translated into practices of recognition, inclusion, and pedagogical care for the body and for the educational relationship with human movement [14,29]. In fact, recent studies in PETE show that students at this stage of teacher education tend to recognize the importance of sustainability, social justice, and professional responsibility, but do not always have sufficiently developed didactic repertoires to operationalize these convictions in everyday teaching [10,29,30].

In contrast, the lowest levels of performance are found in disciplinary and pedagogical content knowledge of the disciplinary curriculum, as well as in the creation of a respectful and well-organized environment, which carries particularly sensitive implications for teacher preparation. This is not simply because these are “more difficult” areas, but because they are dimensions that mediate the possibility of transforming pedagogical principles into real learning experiences and reveal concrete possibilities for transfer to the contexts that teachers actually face [31]. Chilean literature on Initial Teacher Education in PE had already pointed to tensions among ministerial guidelines, teacher education paradigms, and curricular implementation, showing that continuity does not always exist between the discourse of teacher education and the forms of knowledge effectively mobilized by students in teaching contexts [12]. In turn, recent research on learning assessment in Initial Teacher Education in PE indicates that strongly teacher-centered logics persist, with limited student participation and an unstable coexistence between formative approaches and more traditional certification mechanisms [32]. This finding is relevant because difficulties in organizing respectful pedagogical environments while, at the same time, translating disciplinary knowledge into appropriate didactic decisions directly compromise the social sustainability of the educational process, by increasing the risk of exclusion, superficial learning, and the reproduction of inequalities in participation [33,34].

This pattern is consistent with an ongoing international discussion on the gap between the training outlined in Initial Teacher Education curricula and actual pedagogical practice [35]. In the Chilean context, several studies have pointed out that the NDA provides useful information, but also reveals misalignments between the declared curriculum and the effective consolidation of professional competencies during initial teacher education [9,19].

In PE, this gap takes on particular significance because social sustainability is not a peripheral content area, but rather a quality inherent to pedagogical practice itself [36]. PE lies at the intersection of well-being, coexistence, inclusion, and the development of prosocial competencies such as cooperation, teamwork, and respect for diversity; therefore, teaching well does not simply mean organizing motor activity or promoting spaces for sports competition, but rather generating. Conditions for conscious participation, bodily recognition, and relational justice [13,14]. In this sense, the recent literature shows that the integration of sustainability into PE has advanced, but still in an uneven manner, and, in many cases, has been supported more by favorable discourse than by deep methodological transformations [33,37]. Lohmann and Goller observed that teacher educators themselves report. Conceptual ambiguities and practical difficulties when translating education for sustainable development into concrete curricular decisions an obstacle that is particularly significant given that they are the actors expected to lead such curricular design in teacher education programs [38].

Arslan [4] and Xu [5] emphasized that teacher competency assessment systems are useful when they make it possible to monitor progress and guide improvement, but they lose strength when they reduce professional complexity to decontextualized indicators. This is consistent with Biesta’s [7] warning about the risks of an accountability culture that

oversimplifies teaching. In the specific field of sustainability in disciplines other than PE, Blom and Karrow [39] show that the integration of environmental education and sustainability into teacher education remains uneven at the international level. Therefore, the implications for Initial Teacher Education in PE are substantial. First, it seems necessary to strengthen the integration of subject-specific didactics, early practicum experiences, evidence analysis, and systematic feedback mechanisms, as suggested by reviews on formative assessment in PE [34,40]. Second, it is advisable to expand the use of active methodologies with a curricular purpose, since the evidence indicates that these foster reflective learning, commitment to sustainability, and the development of transferable professional competencies [41]. Third, teacher education should more decisively incorporate digital competencies applied to the specific demands of PE, not merely as instrumental innovation. Rather than replacing bodily activity, technologies relevant to this discipline such as wearable devices and applications for monitoring physical activity, heart rate, or step counts; video analysis of motor performance; and platforms for formative feedback and learning documentation can support more flexible, accessible, inclusive, and sustainable learning environments, a potential particularly highlighted in recent reviews [42,43].

The results show a broadly balanced pattern between women and men, with women scoring slightly higher in personal-social development and professional ethics. Overall, the findings suggest comparable pedagogical preparation in PETE and support using the NDA as a diagnostic tool for improving teacher education rather than only for assessment.

Finally, the results reaffirm that the NDA should be understood above all as a diagnostic instrument for educational sustainability, and previous studies on the NDA in Chile have already shown that its legitimacy depends less on its classificatory capacity than on its potential to inform curricular decisions and strengthen improvement processes [9,11]. If it is used in a punitive manner or reduced to an accountability mechanism, assessment may reinforce standardization biases that the critical literature itself has questioned [7,8]. On the other hand, if it is understood as an input for identifying blind spots in teacher education, then the NDA can fulfill a strategically sustainable function and help the teacher education system learn from itself, adjust its curricular trajectories, and strengthen its capacity to support students throughout their training process, ultimately contributing to the preparation of teachers capable of teaching with equity, depth, and public responsibility [2,28]. In this sense, the main contribution of the study is to show that the social sustainability of Initial Teacher Education in PE is not determined solely by adherence to standards, but by the institutional capacity to turn them into consistent and transformative pedagogical practice [2,9,10].

#### *Limitations and Future Research Directions*

This study has several limitations. Its cross-sectional design prevents causal inferences and does not allow competence development to be tracked over time. In addition, it relies on secondary data from a standardized written instrument, which captures mainly quantifiable aspects of pedagogical knowledge and does not fully account for the situated, embodied, and relational competencies central to Physical Education. Achievement was also operationalized as the proportion of correct responses, so qualitative and observational evidence is needed to better relate these findings to teaching practice. Finally, unequal group sizes reduce statistical power in some comparisons. Future research should adopt longitudinal and mixed-methods designs, integrate qualitative approaches, and examine the institutional and curricular factors underlying the weaknesses observed in curricular pedagogical content knowledge and learning environment management.

## 5. Conclusions

In conclusion, the present study made it possible to identify differences in the percentages of achievement among students in Initial Teacher Education in PE across the pedagogical standards assessed by the 2024 NDA, both among the different standards analyzed and according to university, sex, and administrative dependency. Although the results show a generally comparable level of performance within the system, the disaggregated analysis made it possible to identify institutional variations that reflect differentiated training processes. In particular, the highest levels of achievement were concentrated in standards related to professional ethics, student learning and development, and strategies aimed at achieving deep learning, whereas the lowest levels of performance were observed in areas associated with pedagogical content knowledge of the curriculum and the construction of respectful and organized environments. This pattern suggests that teacher education shows relevant progress in normative and interpretive components of the teaching profession, but still reveals training-related bottlenecks in key dimensions for pedagogical mediation and the management of learning in real contexts.

The differences observed according to administrative dependency show a systematic trend in which private institutions belonging to the G9 achieve, across several standards, slightly higher mean scores than public universities and private universities not affiliated with the G9. However, the effect sizes identified are small, so these differences should be interpreted as gradual variations within a common teacher education framework, rather than as substantial gaps between types of institutions. Regarding the analysis by sex, performance showed a balanced trend between women and men, with statistically significant differences being found only in some standards and always with small magnitudes.

Overall, the findings confirm the usefulness of standards-based analysis of the NDA as a tool for making visible specific strengths and challenges in initial teacher education. The disaggregation of results makes it possible to identify areas in which it is necessary to strengthen curricular coherence, the articulation between theory and practice, subject-specific didactics, and feedback mechanisms. In this sense, the NDA should be understood as an instrument oriented toward continuous improvement, capable of providing relevant empirical evidence for processes of institutional self-assessment, curricular redesign, and quality assurance. From the perspective of educational sustainability, its value lies not only in measuring performance, but also in strengthening the capacity of teacher education systems to identify weaknesses, adjust their trajectories, and promote pedagogical responses that are more coherent, equitable, and sustainable over time. In increasingly diversified higher education systems, the disaggregated analysis of pedagogical competencies constitutes a relevant means of understanding more precisely how teacher education performance is configured and how diagnostic assessment policies can contribute not only to quality assurance, but also to the social sustainability of teacher education, fostering programs that are more reflective, inclusive, and capable of responding to the contemporary challenges of school PE.

**Author Contributions:** All authors contributed to the study design, manuscript writing, and final revision. F.G.-F., B.C.-T., L.A.-M. and J.G.-F. were responsible for preparing the initial draft. C.M.-A., J.R.-V. and S.P.-T. contributed to the methodological design and data analysis. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by project Fondecyt Regular No. 1230609 and Fondecyt Regular No. 1241373.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The dataset used in this study is available from the corresponding author upon request and may also be accessed through the following public repository: <https://datosabiertos.mineduc.cl/> (accessed on 11 June 2026).

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

## References

1. Navarrete-Cazales, Z. Modelos y programas de formación de profesores en Chile y México. *Rev. Esp. Educ. Comp.* **2023**, *44*, 165–183. [CrossRef]
2. UNESCO. *Education for Sustainable Development: A Roadmap*; UNESCO: Paris, France, 2020. [CrossRef]
3. Aceska, N. The Role of Teachers' Competencies in Education of Sustainable Development. Available online: <https://meetingorganizer.copernicus.org/EGU23/EGU23-7232.html> (accessed on 7 April 2026).
4. Arslan, M. Monitoring and evaluation of teacher competencies at the international level: A comparative study. *Pedagog. Perspect.* **2025**, *4*, 509–521. [CrossRef]
5. Xu, X. Construction of Evaluation and Quality Assurance Systems in Higher Education. *J. High. Educ. Res.* **2025**, *5*, 554. [CrossRef]
6. Guimarães-Mielo, E.; Sabia, C.P.D.P. Accountability e a formação continuada de professores: Dissonâncias no oferecimento de uma educação integral. *Rev. Bras. Educ. Camp.* **2024**, *9*, 19366. [CrossRef]
7. Biesta, G. Why the form of teaching matters: Defending the integrity of education and of the work of teachers beyond agendas and good intentions. *Rev. Educ.* **2022**, *2022*, 13–33.
8. Morocho-Gamboa, I.V.; Aguagallo-González, A.C. Limitaciones de las evaluaciones estandarizadas y propuestas alternativas del aprendizaje. *Rev. Cienc. Desarro.* **2024**, *24*. [CrossRef]
9. Maturana-Castillo, D.; Castillo-Barrientos, E.; Saba-Henry, F.; Cisternas-Vejar, O.; Matus-León, J. Formación del profesorado: La implementación de la Evaluación Nacional Diagnóstica en el sistema chileno: Un balance sobre su uso. *Profesorado* **2025**, *29*, 1–23. [CrossRef]
10. Ruffinelli, A.; Cisternas, T.; Förster, C.; Donoso, F. La Evaluación Nacional Diagnóstica (END): Percepciones y consideraciones para la Formación Inicial Docente desde estudiantes y directores de carreras. *Rev. Calid. Educ.* **2023**, *59*, 268–298. [CrossRef]
11. Donoso-Reyes, F.; Ruffinelli-Vargas, A. ¿Hacia una Formación Inicial Docente de calidad?: La Evaluación Nacional Diagnóstica en las voces de actores partícipes del proceso de redacción e implementación de la Ley 20.903. *Rev. Estud. Exp. Educ.* **2020**, *19*, 125–147. [CrossRef]
12. Poblete-Valderrama, F.; Illanes Aguilar, L.; Linzmayer Gutiérrez, L.; Cenzano Castillo, L.; Quintana Figueroa, C.; Garrido-Méndez, A.; Rivas Morales, C.; Hetz Rodríguez, K. Initial Teacher Training in Physical Education, paradigms of training from the perception of the trainer of trainers, ministerial guidelines and curricular networks: Case study. *Retos* **2023**, *49*, 552–563. [CrossRef]
13. Jadhav, G.-A. Transforming Lives through Physical Education: Focus on Health, Wellness, and Sustainability. *Int. J. Adv. Appl. Res.* **2025**, *56*, 470–472. [CrossRef]
14. Casey, A.; Kirk, D. *Models-Based Practice in Physical Education*; Routledge: London, UK, 2020.
15. Cloes, M. Preparar ciudadanos físicamente bien educados en Educación Física. Expectativas y prácticas. *Retos* **2017**, *31*, 245–251. [CrossRef]
16. Hang-Diem, T.K.; Thao-Phuong, N. Sustainable Teacher Education and Professional Development. *Vietnam J. Educ.* **2022**, *6*, 1–5. [CrossRef]
17. Sá, P.; Lourenço, M.; Carlos, V. Sustainability Competencies in Higher Education Research: An Analysis of Doctoral Theses in Portugal. *Eur. J. Investig. Health Psychol. Educ.* **2022**, *12*, 387–399. [CrossRef] [PubMed]
18. Ortega-Manjarrez, V.A.; Romero-Manjarrez, E.J.; Manjarrez-Fuentes, N.N. Retos del profesorado en competencias digitales y pedagógicas. *Congr. Caribeño Investig. Educ.* **2025**, *5*, 333–338. [CrossRef]
19. Paidicán-Soto, M.Á.; Arredondo-Herrera, P.A. Revisión de alcance de la formación inicial docente en pedagogía de educación básica en Chile. *Rev. Cient. Cienc. Soc.* **2025**, *7*, 125. [CrossRef]
20. Masoud, M.; Alsayd, A.; Kasheem, M. Global Trends and Challenges in Teacher Training: A Bibliometric Analysis (2019–2024). *Bull. Islam. Res.* **2025**, *3*, 705–732. [CrossRef]
21. Oliveira, H.R.D.; Jesus, D.P.D.; Borges, E.M. As pesquisas sobre formação de professores nas reuniões nacionais da ANPEd: Que espaço há para a avaliação? *Rev. Docência Ens. Sup.* **2019**, *9*, 1–15. [CrossRef]
22. Schwippert, K. Zur Situierung der aktuellen Lehrkräftebildungsforschung: Stand und Perspektiven im Rahmen von internationalen Vergleichsuntersuchungen. *BzL-Beitr. Lehrerinnen-Und Lehrerbildung* **2015**, *33*, 7–21. [CrossRef]
23. Santos, F.A.D.; Zanardini, I.M.S.; Zanardini, J.B.; Marques, H.J. A internacionalização da educação e a construção de sistemas de avaliação de classe mundial para a América Latina. *Rev. Educ. Políticas Debate* **2024**, *13*, 1–15. [CrossRef]
24. Cordray, D.S. Secondary Analysis: Methodology. In *International Encyclopedia of the Social & Behavioral Sciences*; Elsevier: Amsterdam, The Netherlands, 2001; pp. 13777–13780. [CrossRef]

25. Valderrama, F.P. *Formación Inicial Docente en Educación Física: Itinerario Formativo a Través de Prácticas Pedagógicas*; Ediciones UCSC: Concepción, Chile, 2025.
26. De Micheaux, P.L.; Drouilhet, R.; Liquet, B. *The R Software: Fundamentals of Programming and Statistical Analysis*; Springer: New York, NY, USA, 2013; 978p.
27. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Routledge: New York, NY, USA, 2013; 567p. [[CrossRef](#)]
28. Sánchez-Villavicencio, M.F.; Escobar-Gómez, E.; Gutiérrez-Segura, F.A.; Jiménez-Huayama, L.J.; Lujan-Segura, E.; Robles-Villanueva, O.A. *Métodos Estadísticos Aplicados a la Investigación Científica con Software SPSS y XLSTAT*; Editorial Mar Caribe: Colonia, Uruguay, 2025. [[CrossRef](#)]
29. Vidal, S.; Kuckuck, M. Pre-Service Teacher Action Competence in Education for Sustainable Development: A Scoping Review. *Sustainability* **2025**, *17*, 3856. [[CrossRef](#)]
30. Ribeiro-Silva, E.; Amaral-da-Cunha, M.; Batista, P. Educating Teachers for Sustainability and Social Justice: A Service-Learning Project in Physical Education Initial Teacher Education. *Educ. Sci.* **2023**, *13*, 1173. [[CrossRef](#)]
31. Baena-Morales, S.; Merma-Molina, G.; Ferriz-Valero, A. Integrating education for sustainable development in physical education: Fostering critical and systemic thinking. *Int. J. Sustain. High. Educ.* **2023**, *24*, 1915–1931. [[CrossRef](#)]
32. De Longhi, A.L.; Rivarosa, A. Los nuevos estándares para la formación docente: Reflexiones y tensiones. *Rev. Educ. Biol.* **2015**, *18*, 5–10. [[CrossRef](#)]
33. Gallardo-Fuentes, F.; Carter-Thuillier, B.; Peña-Troncoso, S.; Pérez-Norambuena, S.; Gallardo-Fuentes, J. Perceptions of Learning Assessment in Practicum Students vs. Initial Teacher Education Faculty in Chilean Physical Education: A Comparative Study of Two Cohorts. *Educ. Sci.* **2025**, *15*, 459. [[CrossRef](#)]
34. Fröberg, A.; Wiklander, P.; Baena-Morales, S.; Lundvall, S. How to teach about sustainable development in physical education? Examples from the perspectives of certified teachers in Sweden. *Front. Educ.* **2023**, *8*, 1294763. [[CrossRef](#)]
35. Hamzah, N.H.; Samsudin, S.; Hidayat, R. Teacher Competence and Readiness for Classroom-Based Assessment in Physical and Health Education: A Systematic Literature Review. *Int. J. Res. Innov. Soc. Sci.* **2025**, *9*, 3206–3226. [[CrossRef](#)]
36. Venturini, N.d.C.; de Sousa, D.O.; Borges, E.V.; Elias, E.V.C.; Miranda, L.L.d.S.; de Moura, L.R.; da Silva, L.L.; Pereira, T.D.G. Formação inicial de professores e a realidade da sala de aula: Um descompasso histórico. *Cuad. Educ. Desarro.* **2025**, *17*, e8085. [[CrossRef](#)]
37. De la Cruz-Campos, J.-C.; Rodríguez-Jiménez, C.; Martínez-Domingo, J.-A.; Berral-Ortíz, B. Sustainable Experiences applied in the teaching of Physical Education. *Educ. Sport Health Phys. Act.* **2022**, *6*, 208–226. [[CrossRef](#)]
38. Baena-Morales, S.; Fröberg, A.; García-Taibo, O.; Ponce-Gea, A.I. Physical Education Teacher Education Students' Perceptions of Competencies for Sustainable Development in Their Degree Program. *Aust. J. Environ. Educ.* **2025**, *41*, 873–888. [[CrossRef](#)]
39. Lohmann, J.; Goller, A. Physical education teacher educators' subjective theories about sustainability and education for sustainable development. *Int. J. Sustain. High. Educ.* **2023**, *24*, 877–894. [[CrossRef](#)]
40. Blom, R.; Karrow, D.D. Environmental and sustainability education in teacher education research: An international scoping review of the literature. *Int. J. Sustain. High. Educ.* **2024**, *25*, 903–926. [[CrossRef](#)]
41. Herrero-González, D.; López-Pastor, V.M.; Manrique-Arribas, J.C.; Moura, A. Formative and shared assessment: Literature review on the main contributions in physical education and physical education teacher education. *Eur. Phys. Educ. Rev.* **2024**, *30*, 493–510. [[CrossRef](#)]
42. Martínez-Valdivia, E.; Pegalajar-Palomino, M.D.C.; Burgos-Garcia, A. Active methodologies and curricular sustainability in teacher training. *Int. J. Sustain. High. Educ.* **2023**, *24*, 1364–1380. [[CrossRef](#)]
43. Zhang, L.; Yang, C.; Zheng, Y. Digital competence for sustainable education of pre-service teachers: A systematic literature review (2014–2024). *Front. Psychol.* **2026**, *16*, 1710983. [[CrossRef](#)] [[PubMed](#)]

**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.