



VNIVERSITAT E VALÈNCIA

[QÆ] Facultat de Formació
del Professorat

Programa de doctorado en Didácticas específicas (3112)

**La Evaluación en Educación Física en Chile:
narrativas estudiantiles, prácticas docentes y calificaciones**

Presentada por:

Fernando Patricio Muñoz Sepúlveda

Dirigida por:

Dra. Alexandra Valencia Peris

Dr. Daniel Martos García

Dr. Alberto Moreno Doña

Valencia, septiembre 2025

Dra. Alexandra Valencia Peris, profesora Titular de Universidad del Departamento de Didáctica de la Educación Física, Artística y Música de la Universitat de València (España).

Dr. Daniel Martos García, profesor Titular de Universidad del Departamento de Didáctica de la Educación Física, Artística y Música de la Universitat de València (España).

Dr. Alberto Moreno Doña, profesor Titular de Universidad de la Facultad de Medicina-Escuela de Educación Parvularia de la Universidad de Valparaíso (Chile).

CERTIFICAN:

Que la presente memoria, titulada “La Evaluación en Educación Física en Chile: narrativas estudiantiles, prácticas docentes y calificaciones”, corresponde al trabajo realizado bajo su dirección por D. Fernando Patricio Muñoz Sepúlveda, para su presentación como Tesis Doctoral en el Programa de Doctorado en Didácticas Específicas de la Universitat de València.

Y para que así conste, firman el presente certificado en Valencia, a 05 de septiembre de 2025.

Firmado:

Dra. Alexandra Valencia Peris

Dr. Daniel Martos Garcia

Dr. Alberto Moreno Doña

AGRADECIMIENTOS

Un 25 de mayo del año 2020, sin contactos en común, solo con entusiasmo e ideas, escribí un temeroso correo a Sandra para presentarme y proponer una investigación sobre evaluación. Ese mismo día me respondió, y con ello se abrió un mundo. Desde entonces, su guía rigurosa, generosa y visionaria marcó el rumbo de este proyecto. Gracias, Sandra, por darme la oportunidad de aprender, crecer y soñar.

Este camino no ha sido fácil. Con una hija en brazos y otro hijo en camino, la vida familiar y académica se entrelazaron intensamente. Por eso, mi más profundo agradecimiento es para Javiera, mi esposa y compañera de vida, quien ha sido sostén, impulso y refugio en todo momento.

Agradezco también a Dani y Alberto, por su experiencia, sus correcciones certeras y sus miradas críticas, tan necesarias para que este trabajo alcanzara mayor profundidad y sentido.

A mis hijos, Elena y Lorenzo, gracias por su paciencia, por soportar las ausencias y por enseñarme cada día el valor de estar presente. Siempre traté de estar, lo estoy y lo seguiré estando.

A mis estudiantes y colegas, gracias por escuchar con paciencia cada reflexión que compartí en medio del análisis de datos, por acoger mis ideas, por generar ese ambiente de confianza que tanto necesitamos para pensar, construir y concluir.

Finalmente, agradezco a la Universidad Autónoma de Chile, institución que no solo me ha dado la oportunidad de ejercer como profesor, sino que también ha creído firmemente en mi desarrollo académico, brindándome apoyo institucional, tiempos y respaldo económico para llevar adelante este proceso formativo.

Este doctorado ha sido solo el inicio de un camino que espero transforme y aporte. Escribir no es fácil, pero con guía, apoyo e inspiración, es posible. Y deseo, sinceramente, que este esfuerzo genere impacto.

ÍNDICE

RESUMEN	1
ABSTRACT	3
RESUM	5
CAPÍTULO 1. INTRODUCCIÓN	7
1.1. <i>Introducción</i>	8
1.2. <i>Estructura de la tesis y objetivos de la investigación</i>	12
CAPÍTULO 2. MARCO TEÓRICO	15
2.1. <i>Conceptualización</i>	16
2.1.1. Evaluación	16
2.1.2. Evaluación educativa	17
2.1.3. ¿Evaluación o calificación?	20
2.1.4. Estrategia, técnica, procedimiento e instrumento	22
2.1.5. Tipos de evaluación educativa	25
2.2. <i>Paradigmas en la evaluación educativa</i>	28
2.3. <i>Propósitos de la evaluación educativa</i>	33
2.3.1. Evaluación con propósito de diagnóstico	34
2.3.2. Evaluación para rendición de cuentas y mejora institucional	35
2.3.3. Evaluación con fines de orientación y apoyo al aprendizaje	36
2.3.4. Nuevos propósitos de la evaluación en la era digital e inteligencia artificial	37
2.4. <i>Enfoques de evaluación educativa</i>	38
2.4.1. Evaluación formativa y compartida	40
2.4.2. Evaluación auténtica	45
2.4.3. Evaluación por competencias	46
2.4.4. Evaluación digital, inteligente y multimodal	48
2.4.5. Retos y desafíos en la implementación de enfoques	50
2.5. <i>La evaluación en contextos locales</i>	52
2.5.1. Evaluación educativa en América Latina	52
2.5.2. Evaluación educativa en Chile	57

ÍNDICE

2.6. <i>La evaluación en Educación Física</i>	61
2.6.1. La historia de la Educación Física y su relación con la evaluación	61
2.6.2. Paradigmas y enfoques de evaluación en Educación Física	63
2.6.3. La Educación Física en Chile y su relación con la evaluación	65
2.6.4. La formación del profesorado en Educación Física	68
2.7. <i>Estado de la cuestión</i>	70
2.7.1. Investigaciones en evaluación desde la perspectiva del estudiantado.....	73
2.8.2. Investigaciones sobre las concepciones y prácticas del profesorado	75
2.8.3. Desafíos actuales en la investigación sobre evaluación en Educación Física.....	76
CAPÍTULO 3. MARCO METODOLÓGICO.....	79
3.1. <i>Paradigma de investigación</i>	80
3.2. <i>Metodología etapa 1</i>	83
3.2.1. Enfoque de investigación	83
3.2.2. Técnicas e instrumentos de recolección de datos etapa 1	84
3.2.3. Participantes etapa 1	85
3.2.4. Análisis de datos etapa 1	88
3.3. <i>Metodología etapa 2</i>	94
3.3.1. Enfoque de investigación	94
3.3.2. Técnicas e instrumentos de recolección de datos etapa 2	95
3.3.3. Participantes etapa 2.....	98
3.3.4. Análisis de datos etapa 2	99
3.4. <i>Consideraciones éticas de la investigación</i>	106
3.5. <i>Criterios de validez considerados en la investigación</i>	108
3.5.1. Credibilidad.....	109
3.5.2. Transferibilidad	109
3.5.3. Dependibilidad	110
3.5.4. Confirmabilidad	110
3.5.5. Reflexividad	110
CAPÍTULO 4. ESTUDIO 1	112
4.1. <i>Metodología</i>	113

ÍNDICE

4.1.1.	La investigación narrativa	113
4.1.2.	Participantes	114
4.1.3.	Procedimiento.....	115
4.1.4.	Análisis de datos	116
4.2.	<i>Resultados</i>	118
4.3.	<i>Discusión</i>	122
4.4.	<i>Conclusiones estudio 1</i>	126
CAPÍTULO 5. ESTUDIO 2		128
5.1.	<i>Metodología</i>	129
5.1.1.	Participantes	130
5.1.2.	Procedimiento.....	131
5.1.3.	Análisis de datos.....	132
5.2.	<i>Resultados</i>	135
5.3.	<i>Discusión</i>	144
5.4.	<i>Conclusiones estudio 2</i>	147
CAPÍTULO 6. ESTUDIO 3		148
6.1.	<i>Metodología</i>	149
6.1.1.	Participantes	149
6.1.2.	Técnicas de recogida de datos	150
6.1.3.	Cuestiones éticas y de procedimiento	151
6.1.4.	Análisis de datos.....	152
6.2.	<i>Resultados</i>	154
6.3.	<i>Discusión</i>	159
6.4.	<i>Conclusiones estudio 3</i>	163
CAPÍTULO 7. ESTUDIO 4		164
7.1.	<i>Metodología</i>	165
7.1.1.	Fase cualitativa.....	166
7.1.2.	Fase cuantitativa.....	168
7.1.3.	Consideraciones éticas	169
7.2.	<i>Resultados</i>	170

ÍNDICE

7.3. <i>Discusión</i>	176
7.4. <i>Conclusiones estudio 4</i>	179
CAPÍTULO 8. CONCLUSIONES GENERALES.....	181
8.1. <i>Conclusión general de la tesis</i>	182
8.2. <i>Limitaciones y prospectiva</i>	184
Anexo 1. Acta comité de ética	241
Anexo 2. Consentimiento informado estudiantado.....	243
Anexo 3. Consentimiento informado profesorado.....	245
Anexo 4. Instrumento narrativo	247
Anexo 5. Guion de grupo focal para estudiantado.....	250
Anexo 6. Guion de entrevista semiestructurada para el profesorado.....	252
Anexo 7. Registro oficiales calificaciones nacionales desde MINEDUC	254

TABLAS

Tabla 1. Organización de la tesis doctoral con los objetivos del plan de investigación.	14
Tabla 2. Síntesis de clasificación clásica de la evaluación educativa.	25
Tabla 3. Dimensiones comparativas de los paradigmas de evaluación educativa.	32
Tabla 4. Síntesis de normativas asociadas a la evaluación educativa chilena.	57
Tabla 5. Organización metodológica de la tesis según los objetivos y etapas.	82
Tabla 6. Distribución según género de las personas participantes.	86
Tabla 7. Participación Grupos focales.	87
Tabla 8. Matriz de categorías emergentes de los relatos narrativos, desde los contenidos curriculares.	91
Tabla 9. Matriz de categorías emergentes desde los Grupos Focales.	93
Tabla 10. Bases de datos de calificaciones utilizadas en este estudio.	97
Tabla 11. Caracterización del profesorado entrevistado.	98
Tabla 12. Matriz de categorías sobre experiencias del profesorado en evaluación.	100
Tabla 13. Matriz de categorías Conceptualización y prácticas evaluativas.	101
Tabla 14. Caracterización de los instrumentos de evaluación analizados.	102
Tabla 15. Matriz de categorías desde los instrumentos de evaluación.	102
Tabla 16. Matriz de categorías y subcategorías emergentes del análisis de las entrevistas e instrumentos de evaluación.	104
Tabla 17. Distribución según género de las personas participantes.	114
Tabla 18. Matriz de categorías emergentes de los relatos narrativos.	116
Tabla 19. Distribución según género de las personas participantes.	130
Tabla 20. Participación Grupos focales.	131
Tabla 21. Matriz de categorías emergentes de los relatos narrativos, desde los contenidos curriculares.	133
Tabla 22. Matriz de categorías emergentes desde los Grupos Focales.	134
Tabla 23. Caracterización del profesorado entrevistado.	150
Tabla 24. Caracterización de los instrumentos de evaluación analizados.	151

ÍNDICE

Tabla 25. Matriz de categorías y subcategorías emergentes del análisis de las entrevistas e instrumentos de evaluación.	153
Tabla 26. Caracterización del profesorado entrevistado.	166
Tabla 27. Matriz de categorías elaboradas desde las entrevistas.	167
Tabla 28. Bases de datos de calificaciones utilizadas en este estudio.	169
Tabla 29. Calificaciones de la asignatura de EFyS en función del nivel de enseñanza desde el 2018 al 2022.	173
Tabla 30. Calificaciones de la asignatura de Educación Física y Salud en Chile según región geográfica.	174

FIGURAS

Figura 1. Paradigmas en evaluación educativa.....	31
Figura 2. Distribución por disciplina en los grupos focales.	88
Figura 3. Nube de palabras desde los relatos narrativos.	89
Figura 4. Interpretaciones estudiantiles sobre la evaluación en EF: experiencias, sentidos y consecuencias.	92
Figura 5. Mapa conceptual sobre las experiencias en evaluación con test de condición física.	117
Figura 6. Histograma calificaciones nacionales EFyS 2022. (Escala de 1 a 7).....	172

ACRÓNIMOS

CPEIP: Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas

DIA: Diagnóstico Integral de Aprendizajes

EF: Educación Física

EFyC: Evaluación formativa y compartida

EFyS: Educación Física y Salud

EpP: Evaluación para la Praxis

ERCE: Estudio Regional Comparativo y Explicativo

IAG: Inteligencia Artificial generativa

LGE: Ley General de Educación

LOGSE: Ley Orgánica General del Sistema Educativo (España)

MBE: Marco para la Buena Enseñanza

MINEDUC: Ministerio de Educación (Chile)

NASPE: National Association for Sport and Physical Education

OCDE: Organización para la Cooperación y el Desarrollo Económicos

PEF: Pedagogía en Educación Física

PISA: Programme for International Student Assessment

REFyCE: Red de Evaluación Formativa y Compartida en Educación

SAE: Sistema de Admisión Escolar

SIMCE: Sistema de Medición de la Calidad de la Educación

TFG: Trabajo Final de Grado

UTP: Unidad Técnico-Pedagógica

RESUMEN

Investigar la evaluación en EFyS es clave, ya que esta práctica no solo mide aprendizajes, sino que revela concepciones profundas sobre la enseñanza, el rol docente y la relación con el estudiantado. En esta asignatura, históricamente tensionada entre modelos centrados en el rendimiento físico y enfoques orientados al desarrollo integral, la evaluación adquiere una dimensión pedagógica, ética y cultural que merece ser problematizada. En este contexto, la presente tesis doctoral se propuso profundizar en el proceso evaluativo desde las experiencias del estudiantado y las percepciones y prácticas del profesorado en ejercicio, aportando una comprensión situada, crítica y holística del fenómeno evaluativo en EFyS en Chile. La investigación se desarrolló en dos etapas complementarias. La primera se centró en estudiantado universitario de pedagogía, con el objetivo de profundizar sus experiencias escolares sobre la evaluación en EF, se utilizaron relatos narrativos y grupos focales, lo que permitió explorar vivencias subjetivas, significados atribuidos y huellas emocionales del proceso evaluativo. La segunda etapa se enfocó en el profesorado en ejercicio, indagando en sus concepciones y procedimientos evaluativos, junto a la correspondencia entre sus discursos y las calificaciones otorgadas. Esta fase incorporó entrevistas semiestructuradas, análisis de instrumentos de evaluación y un estudio cuantitativo exploratorio mediante bases de datos de calificaciones oficiales del MINEDUC (2018–2022), conformando así un diseño mixto secuencial. Los Estudios 1 y 2 revelan que el estudiantado fue evaluado, principalmente, a través de test de condición física con criterios poco explícitos y sin retroalimentación significativa. Estos procedimientos, vividos con frecuencia como injustos o excluyentes, evocaron emociones intensas como ansiedad, vergüenza o frustración, especialmente en quienes no se ajustaban a los estándares esperados. En muchos casos, estas experiencias derivaron en estrategias de evitación y en una desvinculación con la actividad física, tensionando así los propósitos formativos del currículum nacional. Aunque también se identificaron experiencias positivas cuando la evaluación se vinculó con la participación, la empatía docente o la retroalimentación oportuna, la percepción general fue ambivalente. Además, se evidenciaron prácticas discriminatorias en función del género o la corporalidad, lo que refuerza la necesidad de avanzar hacia enfoques más inclusivos. Los Estudios 3 y 4, centrados en el profesorado, permiten comprender cómo sus prácticas evaluativas actuales están moldeadas por experiencias escolares previas y por su formación inicial. Si bien se

constata una transición desde un enfoque técnico hacia otro que valora la actitud, el esfuerzo y el compromiso del estudiantado, esta resignificación no siempre se traduce en sistemas de evaluación claros, coherentes ni fundamentados. La incorporación de indicadores actitudinales en instrumentos validados por las UTP, como rúbricas, listas de cotejo o escalas de apreciación, representa un intento por diversificar los enfoques evaluativos. No obstante, persisten tensiones vinculadas a exigencias institucionales, culturas escolares tradicionales y una limitada sistematización de estas prácticas. El Estudio 4 complementó esta perspectiva mediante un análisis de las calificaciones oficiales en la asignatura entre 2018 y 2022, evidenciando una distribución altamente concentrada en los tramos superiores, con la nota 7,0 como moda en todos los años analizados. Este fenómeno plantea interrogantes respecto a la correspondencia entre las calificaciones otorgadas y los aprendizajes efectivamente logrados, cuestionando tanto la validez como la función pedagógica de la calificación en EFyS, especialmente cuando esta no se sustenta en procesos formativos sólidos. En conjunto, los cuatro estudios permiten afirmar que la evaluación en EFyS en Chile se encuentra en un punto de inflexión. Si bien se identifican esfuerzos genuinos por parte del profesorado por avanzar hacia prácticas más humanas, inclusivas y formativas, persisten elementos del paradigma tradicional que dificultan su transformación. Esta tensión se manifiesta tanto en las vivencias del estudiantado como en los procedimientos docentes, evidenciando la necesidad de revisar críticamente los fundamentos, métodos y sentidos que sustentan la evaluación escolar. Finalmente, la tesis proyecta líneas de acción orientadas al fortalecimiento de la formación inicial y continua del profesorado, promoviendo experiencias auténticas de evaluación formativa, coevaluación, autoevaluación y retroalimentación dialógica. Asimismo, se propone la necesidad de analizar las implicancias éticas y pedagógicas del uso de tecnologías emergentes, como la IA en contextos evaluativos de la disciplina. Todo esto cobra especial relevancia frente al proceso de actualización curricular que se desarrolla en Chile durante el año 2025, ofreciendo una oportunidad para construir una cultura evaluativa más coherente con los principios formativos, de justicia y equidad que deben orientar la EFyS.

ABSTRACT

Investigating assessment in PEH is essential, as this practice not only measures learning but also reveals deep-rooted conceptions about teaching, the teacher's role, and the relationship with students. In this subject, historically marked by tensions between models focused on physical performance and approaches aimed at holistic development, assessment acquires pedagogical, ethical, and cultural dimensions that must be critically addressed. In this context, the present doctoral thesis aimed to deepen the understanding of assessment processes through the lived experiences of students and the perceptions and practices of in-service teachers, contributing a situated, critical, and holistic perspective of assessment in PEH within the Chilean school system. The research was structured in two complementary stages. The first focused on pedagogy students and explored their past school experiences with assessment in PE. Narrative accounts and focus groups were used to examine subjective experiences, attributed meanings, and emotional traces related to assessment practices. The second stage addressed in-service teachers, exploring their assessment conceptions and procedures, along with the correspondence between their discourse and the grades assigned. This phase included semi-structured interviews, analysis of assessment tools, and an exploratory quantitative study using official grading databases from the Chilean Ministry of Education (MINEDUC) for the period 2018–2022, forming a sequential mixed-methods design. Studies 1 and 2 revealed that students were mainly assessed through physical fitness tests, often lacking clear criteria and meaningful feedback. These procedures were frequently perceived as unfair or exclusionary, evoking intense emotions such as anxiety, shame, or frustration, especially among those who did not meet expected standards. In many cases, these experiences led to avoidance strategies and disengagement from physical activity, thereby undermining the formative aims of the national curriculum. Although some positive experiences were identified when assessment was connected to student participation, teacher empathy, or timely feedback, the overall perception remained ambivalent. Furthermore, discriminatory practices based on gender or body type were evident, reinforcing the need for more inclusive assessment approaches. Studies 3 and 4, focused on teachers, revealed that their current assessment practices are shaped by prior school experiences and their initial teacher education. While there is a clear shift from a technical model to one that values student attitude, effort, and engagement, this redefinition does not always translate into well-

structured, coherent, or theoretically grounded assessment systems. The incorporation of attitudinal indicators into validated tools such as rubrics, checklists, or rating scales represents an attempt to diversify assessment practices. However, tensions remain due to institutional demands, traditional school cultures, and limited systematization of these practices. Study 4 complemented this perspective with an analysis of official grades in PEH from 2018 to 2022, showing a highly concentrated distribution in the upper grade range, with the maximum grade (7.0) being the most frequent each year. This phenomenon raises questions about the alignment between assigned grades and actual learning outcomes, challenging both the validity and pedagogical function of grading in PEH, especially when not supported by solid formative processes. Taken together, the four studies confirm that assessment in Chilean PEH is at a turning point. Although genuine efforts are observed among teachers to adopt more humane, inclusive, and formative practices, elements of the traditional paradigm continue to hinder meaningful transformation. This tension is present in both students' experiences and teachers' procedures, highlighting the need to critically review the foundations, methods, and purposes that underpin school-based assessment. Finally, this thesis proposes several action lines aimed at strengthening initial and continuing teacher education. These include promoting authentic experiences with formative assessment, peer and self-assessment, and dialogic feedback. It also recommends analyzing the ethical and pedagogical implications of using emerging technologies, such as generative artificial intelligence, in assessment contexts. All of this becomes particularly relevant in light of the curricular reform currently underway in Chile during 2025, which offers an opportunity to build an assessment culture that is more aligned with the principles of meaningful learning, justice, and equity that should guide PEH.

RESUM

Investigar l'avaluació en Educació Física i Salut (EFiS) és clau, ja que esta pràctica no sols mesura aprenentatges, sinó que revela concepcions profundes sobre l'ensenyança, el rol docent i la relació amb l'estudiantat. En esta assignatura, històricament tensionada entre models centrats en el rendiment físic i enfocaments orientats al desenrotllament integral, l'avaluació adquirix una dimensió pedagògica, ètica i cultural que mereix ser problematitzada. En este context, la present tesi doctoral es va proposar aprofundir en el procés avaluatiu des de les experiències de l'estudiantat i les percepcions i pràctiques del professorat en exercici, aportant una comprensió situada, crítica i holística del fenomen avaluatiu en EFiS a Xile. La investigació es va desenrotllar en dos etapes complementàries. La primera es va centrar en estudiantat universitari de pedagogia, amb l'objectiu d'aprofundir les seues experiències escolars sobre l'avaluació en EF, es van utilitzar relats narratius i grups focals, la qual cosa va permetre explorar vivències subjectives, significats atribuïts i petjades emocionals del procés avaluatiu. La segona etapa es va enfocar en el professorat en exercici, indagant en les seues concepcions i procediments avaluatius, al costat de la correspondència entre els seus discursos i les qualificacions atorgades. Esta fase va incorporar entrevistes semiestructurades, anàlisi d'instruments d'avaluació i un estudi quantitatiu exploratori mitjançant bases de dades de qualificacions oficials del Ministeri d'Educació de Xile (2018–2022), conformant així un disseny mixt seqüencial. Els Estudis 1 i 2 revelen que l'estudiantat va ser avaluat, principalment, a través de test de condició física amb criteris poc explícits i sense retroalimentació significativa. Estos procediments, viscuts amb freqüència com a injustos o excloents, van evocar emocions intenses com a ansietat, vergonya o frustració, especialment en els qui no s'ajustaven als estàndards esperats. En molts casos, estes experiències van derivar en estratègies d'evitació i en una desvinculació amb l'activitat física, tensionant així els propòsits formatius del currículum nacional. Encara que també es van identificar experiències positives quan l'avaluació es va vincular amb la participació, l'empatia docent o la retroalimentació oportuna, la percepció general va ser ambivalent. A més, es van evidenciar pràctiques discriminatòries en funció del gènere o la corporalitat, la qual cosa reforça la necessitat d'avançar cap a enfocaments més inclusius. Els Estudis 3 i 4, centrats en el professorat, permeten comprendre com les seues pràctiques avaluatives actuals estan modelades per experiències escolars prèvies i per la seua formació inicial. Si bé es

constata una transició des d'un enfocament tècnic cap a un altre que valora l'actitud, l'esforç i el compromís de l'estudiantat, esta resignificació no sempre es tradueix en sistemes d'avaluació clars, coherents ni fonamentats. La incorporació d'indicadors actitudinals en instruments validats per la Unitat Tècnica Pedagògica, com a rúbriques, llistes d'acarament o escales d'apreciació, representa un intent per diversificar els enfocaments avaluatius. No obstant això, persisteixen tensions vinculades a exigències institucionals, cultures escolars tradicionals i una limitada sistematització d'estes pràctiques. L'Estudi 4 va complementar esta perspectiva mitjançant una anàlisi de les qualificacions oficials en l'assignatura entre 2018 i 2022, evidenciant una distribució altament concentrada en els trams superiors, amb la nota 7,0 com a moda en tots els anys analitzats. Este fenomen planteja interrogants respecte a la correspondència entre les qualificacions atorgades i els aprenentatges efectivament reeixits, qüestionant tant la validesa com la funció pedagògica de la qualificació en EFiS, especialment quan esta no se sustenta en processos formatius sòlids. En conjunt, els quatre estudis permeten afirmar que l'avaluació en EFiS a Xile es troba en un punt d'inflexió. Si bé s'identifiquen esforços genuïns per part del professorat per avançar cap a pràctiques més humanes, inclusives i formatives, persisteixen elements del paradigma tradicional que dificulten la seua transformació. Esta tensió es manifesta tant en les vivències de l'estudiantat com en els procediments docents, evidenciant la necessitat de revisar críticament els fonaments, mètodes i sentits que sustenten l'avaluació escolar. Finalment, la tesi projecta línies d'acció orientades a l'enfortiment de la formació inicial i contínua del professorat, promovent experiències autèntiques d'avaluació formativa, coavaluació, autoavaluació i retroalimentació dialògica. Així mateix, es proposa la necessitat d'analitzar les *implicancias ètiques i pedagògiques de l'ús de tecnologies emergents, com la Intel·ligència artificial generativa en contextos avaluatius de la disciplina. Tot això cobra especial rellevància enfront del procés d'actualització curricular que es desenrotlla a Xile durant l'any 2025, oferint una oportunitat per a construir una cultura avaluativa més coherent amb els principis formatius, de justícia i equitat que han d'orientar la EFiS.



VNIVERSITAT DE VALÈNCIA

CAPÍTULO 1

INTRODUCCIÓN

1.1. Introducción

Esta tesis doctoral se centra en el estudio de la evaluación en la asignatura de EFyS en el contexto chileno, con el propósito de profundizar en las experiencias del estudiantado y analizar las concepciones y prácticas evaluativas que el profesorado en ejercicio despliega en sus espacios pedagógicos. La investigación asume que comprender los sentidos que adquiere la evaluación en esta asignatura requiere atender tanto a las vivencias del estudiantado como a las decisiones metodológicas y valorativas del profesorado, enmarcadas en trayectorias formativas, marcos normativos y culturas escolares particulares. A través de un diseño que articula diversas técnicas y fuentes de información, se busca generar conocimiento situado sobre los modos en que se evalúa en EFyS, considerando sus tensiones, continuidades y transformaciones en el tiempo.

¿Por qué investigar la evaluación? La evaluación es un proceso inherente al ser humano que se manifiesta en las diversas esferas de la vida cotidiana (Madaus et al., 2012; Sandoval et al., 2022). Esta afirmación se basa en la premisa de que la capacidad de evaluar es fundamental para la toma de decisiones y un componente esencial del juicio y la reflexión humana (Facione, 2007; Tyler, 1950). En el ámbito educativo, esta función cobra especial relevancia, no solo por su impacto en el aprendizaje sino, también, por lo que revela sobre las concepciones pedagógicas, éticas y humanas de quien evalúa. Tal como afirma Santos Guerra (2003) en el título de su artículo publicado en *Enfoques Educativos*, “Dime cómo evalúas y te diré qué tipo de profesional y de persona eres” (p. 69), la evaluación no es una práctica neutra, sino un acto profundamente ligado a la visión del mundo, los valores y las intenciones de cada educador.

Esta investigación se inscribe en un giro paradigmático de alcance internacional que impulsa el tránsito desde modelos tradicionales de evaluación, centrados en la medición y la calificación, hacia enfoques orientados al aprendizaje profundo, la participación activa del estudiantado y el juicio ético del profesorado. Este cambio ha sido respaldado por marcos como el Espacio Europeo de Educación Superior, consolidado a partir de la Declaración de Bolonia (European Ministers of Education, 1999), que promueve una formación basada en competencias, el uso de metodologías activas y una evaluación alineada con los procesos

formativos. En este contexto, destaca el trabajo de la REFyC activa desde 2005, la cual ha desarrollado experiencias de innovación e investigación en torno a la evaluación formativa en todos los niveles educativos (REFyC, s.f.). Entre sus aportes recientes se encuentra la tesis doctoral de Atienza (2022), centrada en las percepciones del estudiantado de formación inicial en EF sobre un modelo de evaluación formativa implementado en la universidad. En Chile, esta orientación adquiere relevancia frente a transformaciones normativas como la Ley N.º 20.903, el Decreto N.º 67 (2018) y los Estándares del CPEIP (2021), que promueven una evaluación centrada en la mejora del aprendizaje. En este escenario, el presente estudio se inscribe en una línea investigativa crítica y cualitativa que busca contribuir a la transformación de las prácticas evaluativas en la asignatura de EFyS.

El concepto de evaluación admite múltiples interpretaciones, que varían según los enfoques pedagógicos, contextos institucionales y finalidades que se le atribuyan. Desde una perspectiva técnica o tradicional, ha sido históricamente asociada a funciones de control, medición de resultados, validación de objetivos y rendición de cuentas, con un fuerte énfasis en la estandarización y la cuantificación del aprendizaje (Morales y Vargas, 2004). Sin embargo, es imprescindible reconocer que la evaluación constituye un componente esencial de todo proceso educativo, y que su valor radica en la capacidad de acompañar, orientar y mejorar el aprendizaje (López-Pastor et al., 2006; Prieto, 2015). Diversos autores coinciden en que su propósito no debe limitarse a clasificar o etiquetar al estudiantado, sino a favorecer procesos reflexivos, formativos y significativos (Herrera, 2018; Moreno, 2016; Talanquer, 2015). Esta premisa también se aplica a la asignatura de EFyS, cuyo objetivo es contribuir al logro de aprendizajes integrales mediante el desarrollo de competencias motrices, sociales, cognitivas y actitudinales (Ministerio de Educación, 2015).

Según Moreno-Doña et al. (2013), la evaluación cumple un rol esencial en el desarrollo de la motricidad humana al fomentar estilos de vida activos en la población escolar, lo que se encuentra alineado con los propósitos establecidos en las Bases Curriculares de la asignatura de EFyS. En el caso chileno, la evaluación en esta asignatura se encuentra regulada por el Decreto Exento N.º 67 (Ministerio de Educación, 2018), el cual establece orientaciones sobre evaluación, calificación y promoción escolar. Esta normativa plantea que las

calificaciones deben reflejar el nivel de logro de los objetivos de aprendizaje y permite a los establecimientos educacionales diseñar sus propios reglamentos de evaluación dentro del marco establecido, lo que genera diferencias relevantes en los procesos evaluativos entre centros escolares.

El decreto mencionado articula principios que apuntan a fortalecer una evaluación formativa, flexible y centrada en el proceso; sin embargo, su implementación ha sido diversa y, en muchos casos, sigue predominando una lógica de evaluación sumativa (Rosales, 2014). Esta situación tensiona la finalidad pedagógica de la evaluación, ya que, si bien se promueve una cultura más reflexiva y centrada en el aprendizaje, en la práctica persisten modelos tradicionales que privilegian la calificación como instrumento de control. Frente a esto, diversos autores sostienen que la evaluación debe concebirse como un proceso que favorezca el aprendizaje, más allá de su función certificadora (Atienza et al., 2018; Gómez et al., 2006; López-Pastor et al., 2006). Bajo esta mirada, se entiende la evaluación como un acto pedagógico, ético y transformador, orientado al diálogo, la comprensión y la mejora (Santos Guerra, 1995), el cual debe ser analizado en su complejidad contextual y relacional (Collier, 2011; Moreno-Doña y Poblete-Gálvez, 2015).

En este escenario surgen interrogantes esenciales que orientan el presente trabajo: ¿cómo se desarrolla la evaluación en la asignatura de EFyS?, ¿quiénes participan en este proceso y con qué propósito?, ¿qué instrumentos y criterios se utilizan?, ¿cuándo y por qué se evalúa? Estas preguntas permiten trascender una mirada meramente técnica de la evaluación, invitando a considerar los significados que construyen los actores educativos en torno a sus prácticas, así como la dimensión cultural, temporal y pedagógica que las configura (Peña-Troncoso et al., 2019).

Por una parte, las competencias necesarias para que el profesorado de EFyS pueda evaluar los aprendizajes escolares se construyen, en gran medida, durante la formación inicial, mediante asignaturas específicas de la disciplina (Muñoz et al., 2017). No obstante, resulta pertinente cuestionar cómo las experiencias evaluativas vividas previamente, tanto en su etapa escolar como universitaria, influyen en la forma en que las y los docentes conciben y ejercen la evaluación. Tanto la construcción de la identidad profesional, como los saberes

pedagógicos que se ponen en juego en la enseñanza, están profundamente marcados por dichas experiencias, creencias y relaciones (Merellano-Navarro et al., 2019).

Desde la perspectiva del estudiantado también emergen preguntas relevantes: ¿cómo son vividas las experiencias de evaluación en EFyS?, ¿de qué manera estas inciden en su percepción de la asignatura y en su actitud hacia la práctica físico-motriz? La evaluación, en tanto proceso relacional y cargado de sentido, puede actuar como un elemento motivador o, por el contrario, como una barrera para la participación, la adherencia y la valoración de la asignatura. Investigaciones recientes han evidenciado el impacto que tienen estas experiencias evaluativas en la motivación y percepción del alumnado (Elif y Zekiye, 2020; Navarro-Patón et al., 2020).

Considerando lo anterior, esta tesis doctoral se articula en torno a la necesidad de comprender de forma situada, crítica y reflexiva el fenómeno de la evaluación en EF. A partir de ello, se establecen los objetivos que orientan esta investigación, los cuales se presentan en el apartado siguiente.

1.2. Estructura de la tesis y objetivos de la investigación

El propósito de esta tesis es profundizar en el proceso de evaluación en la asignatura de EFyS mediante el análisis de las experiencias, percepciones y prácticas de quienes enseñan y aprenden en este campo. Para ello, se optó por una estrategia metodológica que articula dos etapas complementarias: una centrada en el estudiantado universitario de pedagogía, y otra en el profesorado de EFyS en ejercicio. Esta decisión responde a la necesidad de comprender el fenómeno evaluativo desde una perspectiva amplia y situada, reconociendo que tanto el profesorado como el estudiantado son agentes fundamentales en la configuración y reproducción de las prácticas evaluativas (Gallardo-Fuentes et al., 2022; López-Pastor et al., 2006).

Numerosos estudios han demostrado que las experiencias previas de evaluación vividas por el profesorado, tanto en su formación inicial como en su etapa escolar, influyen significativamente en sus concepciones y formas de evaluar (Atienza, 2022; Muñoz-Sepúlveda et al., 2025). Al mismo tiempo, investigaciones recientes han subrayado el impacto que tienen estas prácticas evaluativas en la percepción, motivación y participación del estudiantado en las clases de EF (Elif y Zekiye, 2020; Navarro-Patón et al., 2020). Por tanto, atender a ambas perspectivas resulta clave para construir una visión holística del proceso evaluativo y para identificar posibles tensiones, continuidades y oportunidades de transformación en los contextos escolares y universitarios (Moreno-Doña y Poblete-Gálvez, 2015; Sonlleve et al., 2018;).

Desde esta lógica, la tesis no solo busca describir prácticas, sino también comprender los sentidos que adquiere la evaluación en EFyS desde las voces y trayectorias de quienes han sido evaluados y de quienes evalúan. Esta doble aproximación permite generar conocimientos relevantes para fortalecer los procesos de formación docente y para la mejora de las prácticas evaluativas en el sistema escolar chileno, en línea con los marcos normativos y pedagógicos vigentes.

Los objetivos específicos de cada etapa de la investigación se describen a continuación:

Etapla 1: Para esta parte de la investigación se formuló una propuesta en base a textos narrativos, experiencias ligadas a la evaluación en EF en edad escolar, por lo que se requirieron personas que hubiesen finalizado toda su etapa escolar. Para ello, se involucró específicamente a estudiantes universitarios de una facultad de Educación, con los siguientes objetivos específicos:

- Caracterizar las experiencias vividas sobre evaluación en las clases de EF escolar en estudiantes de una Facultad de Educación.
- Indagar sobre la influencia del proceso evaluativo de la EF escolar en la percepción y actitud hacia la asignatura de los estudiantes de una Facultad de Educación.

Etapla 2: Esta fase consideró la participación del profesorado de la asignatura de EFyS en ejercicio, con los siguientes objetivos específicos:

- Profundizar en la conceptualización y perspectivas evaluativas del profesorado de EFyS.
- Analizar las prácticas evaluativas del profesorado en ejercicio, desde sus experiencias escolares y universitarias en evaluación.
- Analizar las prácticas evaluativas del profesorado de EFyS, considerando tanto sus percepciones como el comportamiento de las calificaciones oficiales entre los años 2018 y 2022.

Esta tesis doctoral se estructura en torno a una secuencia de cuatro estudios que responden de manera articulada a los objetivos generales y específicos del plan de investigación. Para otorgar coherencia y solidez al trabajo, en primer lugar, se plantea un marco teórico común, donde se abordan conceptos fundamentales relacionados con la evaluación educativa, incluyendo su conceptualización, los distintos paradigmas, enfoques, finalidades y su desarrollo en el contexto latinoamericano y chileno.

A continuación, se presenta un marco metodológico general, que explica y fundamenta el diseño investigativo adoptado, así como las decisiones metodológicas que sustentan la construcción del proceso indagativo. Posteriormente, se desarrollan los cuatro estudios, cada

INTRODUCCIÓN

uno de los cuales contiene su propio título, un resumen introductorio, la descripción de la metodología específica empleada, los resultados obtenidos, su discusión respectiva y una conclusión particular. Esta estructura permite un análisis profundo y situado de cada fase del proceso de investigación. En forma de síntesis, la tabla 1, compila las etapas de la investigación, los estudios que emergieron de cada etapa y su relación con los objetivos de investigación.

Tabla 1. Organización de la tesis doctoral con los objetivos del plan de investigación.

Etapas	Estudio	Objetivos de la tesis abordados
Etapas 1. Desde las experiencias y perspectivas del estudiantado.	1. Dime cuánto corres y te diré qué nota mereces. Evaluación en la Educación Física chilena.	Caracterizar las experiencias vividas sobre evaluación en las clases de EF escolar en estudiantes de una Facultad de Educación.
	2. Evaluación en Educación Física escolar y su impacto en la percepción de la asignatura: vivencias narrativas de estudiantes de pedagogía.	Indagar sobre la influencia del proceso evaluativo de la EF escolar en las actitudes del alumnado hacia la asignatura.
Etapas 2. Desde las percepciones y prácticas evaluativas del profesorado de EF.	3. Del rendimiento al esfuerzo y actitud: Enfoques evaluativos en Educación Física desde la experiencia docente.	Profundizar en la conceptualización y perspectivas evaluativas del profesorado de EFyS en ejercicio. Analizar las prácticas evaluativas del profesorado en ejercicio, desde sus experiencias escolares y universitarias en evaluación.
	4. La nota máxima está de moda. Análisis de las prácticas evaluativas en la asignatura de Educación Física y Salud chilena.	Analizar las prácticas evaluativas del profesorado de EFyS en ejercicio, considerando tanto sus percepciones como el comportamiento de las calificaciones oficiales entre los años 2018 y 2022.

Una vez abordados los cuatro estudios, se presenta un capítulo de conclusiones generales, donde se integran los hallazgos de cada uno de ellos para construir una reflexión global sobre el fenómeno investigado. Finalmente, se incluyen los apartados correspondientes a las limitaciones y proyecciones de la investigación, seguidos por los anexos y el marco de referencias general que sustenta teórica y empíricamente el conjunto del trabajo.



VNIVERSITAT DE VALÈNCIA

CAPÍTULO 2

MARCO TEÓRICO

2.1. Conceptualización

2.1.1. Evaluación

El objeto de estudio de esta tesis doctoral es la evaluación, con énfasis particular en su manifestación dentro de la asignatura de EF. Para abordar este fenómeno desde una perspectiva comprehensiva, resulta imprescindible delimitar conceptualmente el término. Etimológicamente, evaluación proviene del francés *évaluation*, derivado del verbo *évaluer*, formado por el prefijo *é-* (del latín *ex-*, que significa “fuera de”) y *valeur* (del latín *valere*, “tener valor o fuerza”) (Harper, 2001). Esta raíz remite a la acción de atribuir o reconocer valor, un sentido que se mantiene en su uso pedagógico. De acuerdo con la Real Academia Española (2014), evaluar significa “señalar el valor de algo” o “estimar los conocimientos, aptitudes y rendimiento de los alumnos”, especialmente en el contexto del sistema educativo formal. En consecuencia, evaluar implica más que medir o calificar; supone emitir un juicio fundamentado sobre el proceso o resultado de un aprendizaje, en función de criterios previamente establecidos.

No obstante, la acción de evaluar trasciende ampliamente el ámbito escolar. En la psicología, por ejemplo, se emplea para valorar variables cognitivas, afectivas y conductuales en procesos diagnósticos e intervenciones (Guba y Lincoln, 1989). En las ciencias sociales, la evaluación permite analizar el impacto, eficacia o pertinencia de políticas públicas, programas sociales y proyectos institucionales, generando información clave para la toma de decisiones (Stufflebeam y Shinkfield, 2007). En el mundo organizacional, se vincula con el desempeño, la eficiencia y la planificación estratégica, como parte de los sistemas de gestión de calidad (Fitzpatrick et al., 2011). Estas diversas aplicaciones muestran que la evaluación constituye una práctica transversal orientada a emitir juicios informados y fundamentados, con el objetivo de comprender, mejorar o reorientar procesos de distinta naturaleza.

Desde mediados del siglo XX, la evaluación comenzó a consolidarse como un objeto de estudio autónomo y transversal en diversas disciplinas. En el ámbito educativo, esta consolidación teórica se inició con modelos ligados a la medición del rendimiento, influenciados por una perspectiva cuantitativa y estandarizadora del aprendizaje (Madaus et al., 2012). Bajo este enfoque, evaluar se concebía como un proceso técnico y neutral destinado a clasificar y verificar logros de aprendizaje mediante pruebas objetivas. No

obstante, a partir de las críticas surgidas en las décadas de 1970 y 1980, comenzaron a desarrollarse enfoques más interpretativos, comprensivos y contextuales, que desplazaron el énfasis desde la medición hacia la mejora continua, la participación y el aprendizaje significativo (Stufflebeam y Shinkfield, 2007).

Así, la evaluación ha sido progresivamente resignificada como una práctica pedagógica que no solo mide, sino que orienta, acompaña y transforma. Se trata de una herramienta que, lejos de limitarse a certificar resultados, se integra de manera compleja en las dinámicas educativas, permitiendo retroalimentar los procesos, involucrar a los actores y contribuir al desarrollo integral de quienes aprenden (López-Pastor et al., 2006; Santos Guerra, 2014). Esta concepción más holística y formativa de la evaluación es la que orienta el marco conceptual de esta investigación.

2.1.2. Evaluación educativa

En el campo educativo, la evaluación ha sido objeto de atención sistemática desde mediados del siglo XX, especialmente a partir de los aportes de Ralph Tyler, quien la definió como un medio sistemático para determinar el grado en que se alcanzan los objetivos instruccionales, estableciendo así una relación directa entre evaluación y control curricular (Tyler, 2009). Posteriormente, Bloom et al. (1971) profundizaron esta perspectiva a través de su influyente taxonomía de objetivos cognitivos, proponiendo una estructura jerárquica del aprendizaje que permitía orientar la planificación y la evaluación de los procesos educativos. A su vez, Scriven (1967) introdujo la distinción entre evaluación formativa y sumativa, marcando un hito teórico en la comprensión de los diferentes propósitos de la evaluación.

Estos enfoques, fuertemente influenciados por los paradigmas conductista y positivista, promovieron una concepción técnica de la evaluación, centrada en la medición objetiva de resultados, la validez cuantitativa de los instrumentos y la estandarización de criterios. Bajo esta mirada, el aprendizaje se concebía como un producto observable y mensurable, y la función principal de la evaluación era verificar su nivel de logro.

No obstante, a partir de la década de 1970, comenzaron a emerger perspectivas críticas que ampliaron y problematizaron esta visión. Paulo Freire propuso una concepción de la evaluación como acto dialógico, ético y liberador, en el cual el conocimiento se construye

mediante la interacción reflexiva entre educador y educando, reconociendo la historicidad, subjetividad y politicidad del acto educativo. En una línea similar, Santos Guerra (1995) señaló que la evaluación no debe entenderse como un simple mecanismo de control, sino como un proceso orientado a la comprensión y a la mejora, profundamente vinculado a las intenciones pedagógicas y a los valores éticos del profesorado. Desde esta perspectiva, evaluar implica asumir una postura ética sobre qué se valora y con qué fines, ya que toda acción evaluativa refleja una determinada concepción de ser humano y de educación.

Complementariamente, Perrenoud (1998) enfatizó la necesidad de avanzar hacia una evaluación más participativa, que incluya prácticas como la autoevaluación y la coevaluación, entendidas como medios para democratizar la enseñanza, fortalecer la autonomía del estudiantado y contribuir a una cultura de justicia educativa. Estos aportes marcaron un viraje hacia enfoques interpretativos y sociocríticos, que reconocen el carácter situado, subjetivo y relacional de la evaluación educativa.

En la actualidad, se promueve una concepción más holística, que no se limita a certificar aprendizajes, sino que los acompaña, orienta y transforma. La evaluación se entiende como un proceso pedagógico integral que favorece la toma de decisiones, el desarrollo de prácticas inclusivas y la reflexión crítica de docentes y estudiantes (Conde-Rojas et al., 2022; Córdoba y Lanuza, 2021). Esta visión reconoce que evaluar no es solo aplicar instrumentos, sino diseñar experiencias significativas de aprendizaje, donde lo afectivo, lo cognitivo y lo contextual se articulan en una lógica formativa.

Autores como Torrance (2012) y Allal (2013) han sostenido que una evaluación verdaderamente educativa debe incorporar no solo el análisis de productos, sino también de los procesos de aprendizaje, considerando dimensiones emocionales, sociales, éticas y culturales que inciden en la experiencia escolar. De este modo, la evaluación se reconfigura como una práctica social situada, en la que se producen sentidos, se configuran relaciones de poder y se establecen criterios de legitimación del saber.

En este mismo sentido, Broadfoot (2007) plantea que la evaluación no se limita a reflejar la realidad educativa, sino que la constituye activamente: produce efectos, genera jerarquías y moldea las identidades de quienes participan en ella. Esta capacidad performativa ha sido

abordada también por Gipps (1999) y Biesta (2009), quienes destacan que la evaluación no solo genera consecuencias individuales, sino también sociales, al definir qué aprendizajes son valorados y cuáles son marginados, y al establecer qué estudiantes son reconocidos como competentes y cuáles no.

Por tanto, pensar la evaluación educativa desde una perspectiva crítica implica asumir su carácter estructurante, ético y performativo. Evaluar no es un acto neutro ni meramente técnico, sino una práctica que incide directamente en las trayectorias escolares, en la construcción de subjetividades y en la distribución de oportunidades de aprendizaje (Biesta, 2009; Gipps, 1999). Esta mirada demanda una reflexión profunda sobre los fines de la evaluación, los marcos que la orientan y los métodos que se utilizan, reconociendo su potencial para reproducir desigualdades o, por el contrario, transformarlas a través de una pedagogía más justa, inclusiva y situada (Broadfoot, 2007; Torrance, 2012).

En los últimos años, la literatura especializada ha insistido en ampliar la comprensión de la evaluación educativa más allá de su dimensión técnica o certificadora, recuperando su carácter formativo, ético y transformador. Según Conde-Rojas et al. (2022), la evaluación debe entenderse como un proceso continuo, que comienza desde el inicio de la enseñanza y se extiende durante todo el aprendizaje, generando información valiosa para la toma de decisiones pedagógicas. Esta perspectiva contrasta con visiones reduccionistas centradas en la calificación, y promueve prácticas de retroalimentación reflexiva y contextualizada. De modo similar, Córdoba y Lanuza (2021) subrayan la importancia de concebir la evaluación como un proceso crítico de análisis de resultados académicos, orientado a la mejora continua más que a la mera acreditación. En este mismo sentido, Cruzado (2022) resalta la centralidad de la evaluación formativa como herramienta para favorecer aprendizajes significativos y promover una participación activa del estudiantado, destacando el rol de la retroalimentación como componente clave. Desde un enfoque antropológico-pedagógico, Sandoval et al. (2022) enfatizan que evaluar es una práctica intrínseca al ser humano, y que en educación implica una reflexión ética sobre el sentido profesional de las decisiones que se toman. Finalmente, Prieto-Castillo et al. (2023) propone repensar la evaluación como una oportunidad para acompañar el aprendizaje con criterios de personalización, colaboración y respeto, alertando sobre los riesgos de emplearla como ejercicio de poder alejado del ideal

pedagógico. En conjunto, estas miradas enriquecen la conceptualización de la evaluación educativa, situándola como una práctica compleja, situada y políticamente significativa.

En consecuencia, se hace indispensable clarificar conceptualmente los elementos que componen el proceso evaluativo, distinguiendo entre evaluación y calificación, además de diferenciar sobre estrategia, técnica, procedimiento e instrumento. Esta distinción no solo permite una mayor rigurosidad conceptual, sino que facilita la comprensión y análisis de las prácticas evaluativas concretas que se implementan en el ámbito educativo. Los siguientes apartados se dedican a desarrollar estas categorías, con el propósito de delimitar sus funciones, alcances y articulaciones dentro del entramado pedagógico de la evaluación. Sin embargo, inmersos en aclaraciones terminológicas, merece la pena, antes que eso, definir la evaluación en contraposición a la calificación.

2.1.3. ¿Evaluación o calificación?

En el marco de esta tesis, resulta fundamental diferenciar conceptualmente los términos evaluación y calificación, ya que suelen confundirse o utilizarse como sinónimos en el discurso educativo cotidiano. Mientras que la evaluación ha sido abordada previamente como un proceso pedagógico de retroalimentación, comprensión y mejora del aprendizaje (López-Pastor, 2006; Prieto, 2015; Santos Guerra, 1995), la calificación representa una operación distinta y de carácter sumativo: es la traducción numérica, simbólica o conceptual del juicio valorativo sobre el rendimiento del estudiante en relación con los objetivos establecidos.

Etimológicamente, el término calificación proviene del latín *qualificare*, que significa atribuir una cualidad o condición a algo o alguien. En el ámbito educativo, esta práctica se formalizó históricamente como una forma de ordenar, comparar y certificar el nivel de logro de los aprendizajes, consolidándose con el desarrollo de los sistemas escolares modernos (López-Pastor, 2005; Orozco-Jutorán, 2006; Sandoval et al., 2022). Su función ha estado tradicionalmente vinculada a la evaluación sumativa, orientada a certificar los aprendizajes alcanzados y a tomar decisiones como la promoción o la acreditación de competencias (Bourdieu y Passeron, 2007; Talanquer, 2015).

Desde el plano normativo, el Decreto N.º 67 del Ministerio de Educación de Chile (2018) define la calificación como “la representación del logro en el aprendizaje a través de un

proceso de evaluación, que permite transmitir un significado compartido respecto a dicho aprendizaje mediante un número, símbolo o concepto”. Este mismo decreto establece una escala numérica de 1.0 a 7.0, siendo 4.0 la nota mínima de aprobación. Asimismo, se otorga autonomía a cada establecimiento educacional para redactar su propio reglamento de evaluación y calificación, siempre que respete los lineamientos generales, lo que introduce márgenes de flexibilidad, pero también de disparidad a nivel nacional.

Este escenario plantea desafíos relevantes, especialmente en un país geográficamente extenso y socioculturalmente diverso como Chile. Hasta la fecha, no se han identificado estudios que analicen de forma sistemática cómo varían los criterios de calificación entre regiones, tipos de escuelas o niveles educativos. La posibilidad de que se estén consolidando culturas evaluativas diferenciadas a lo largo del país constituye una hipótesis válida, aún poco explorada, que merece atención académica.

En el caso específico de la asignatura de EFyS, la calificación suele derivar de criterios heterogéneos que combinan el rendimiento físico, la actitud, la participación o el cumplimiento de tareas. Esta multiplicidad de criterios, sumada a la falta de sistematización y transparencia, ha sido cuestionada por reproducir lógicas propias del paradigma técnico y poco pedagógicas, con escasa vinculación con los objetivos curriculares y los principios de una evaluación para el aprendizaje (Castillo-Retamal et al., 2024; Oliver-Álvarez y Martos-García, 2023).

Estudiar la calificación desde una perspectiva crítica permite problematizar su rol no solo como un instrumento técnico, sino como una práctica social cargada de implicancias pedagógicas, éticas y políticas. Las calificaciones, en tanto dispositivos de selección, comparación y legitimación, configuran expectativas, jerarquías y trayectorias escolares (Bourdieu y Passeron, 2007, p. 183). Así, su análisis debe situarse en el debate más amplio sobre justicia educativa, equidad y democratización del conocimiento, especialmente en asignaturas como EF, donde históricamente han prevalecido prácticas evaluativas excluyentes y poco formativas.

2.1.4. Estrategia, técnica, procedimiento e instrumento

En el transcurso de esta tesis doctoral, se emplearán conceptos relacionados con la evaluación educativa que, debido a su cercanía semántica, suelen prestarse a confusión. Por ello, resulta pertinente delimitar conceptualmente los términos estrategia, técnica, procedimiento e instrumento, con el fin de clarificar sus alcances y relaciones dentro de los procesos de enseñanza-aprendizaje.

- *Estrategia*: Según la Real Academia Española de la Lengua (2014), una estrategia es un “conjunto de acciones planificadas sistemáticamente en el tiempo que se llevan a cabo para lograr un determinado fin”. En el ámbito educativo, este término adquiere un matiz pedagógico, ya que se refiere a un plan flexible e intencionado que orienta las decisiones docentes para facilitar el aprendizaje. En el campo de la evaluación, una estrategia implica coordinar coherentemente métodos, técnicas e instrumentos que permitan valorar el proceso de enseñanza-aprendizaje de forma pertinente y formativa. Según Bixio (2006), las estrategias evaluativas deben responder a una lógica de coherencia pedagógica y contextualización, considerando tanto los objetivos curriculares como las características del estudiantado. Así, planificar la evaluación desde una perspectiva estratégica contribuye a generar prácticas más reflexivas, inclusivas y ajustadas a las realidades escolares. Además, Sandoval Rubilar et al. (2022) destacan que una estrategia evaluativa eficaz debe considerar no solo los aprendizajes esperados, sino también los valores éticos y las características del grupo de estudiantes, integrando así una visión pedagógica situada.

- *Técnica*: De acuerdo con la Real Academia Española de la Lengua (2014), una técnica es un “conjunto de procedimientos y recursos de que se sirve una ciencia o arte”. En el ámbito educativo, y más específicamente en la evaluación, la técnica representa una acción específica dentro de un plan estratégico mayor. Se trata de un medio operativo que define cómo se obtendrá la información para valorar los aprendizajes del estudiantado, mediante actividades concretas, observables y reproducibles. Ejemplos de técnicas incluyen la observación sistemática, la entrevista, el análisis de portafolios, la prueba escrita o la exposición oral. Estas se eligen según el tipo de aprendizaje a evaluar (cognitivo, procedimental o actitudinal) y según el enfoque pedagógico adoptado (formativo, sumativo,

participativo, entre otros) (Zabalza, 2000). Según Cruzado (2022), la técnica seleccionada condiciona la naturaleza de la interacción entre docente y estudiante, así como el tipo de evidencias generadas para sustentar una evaluación justa y significativa. Por ello, su elección debe estar alineada con los principios de equidad, inclusión y coherencia curricular.

- *Procedimiento*: Se concibe como el conjunto ordenado de pasos que permiten llevar a cabo una técnica evaluativa en coherencia con una estrategia didáctica determinada. Representa la dimensión operativa de la evaluación, en tanto define el “cómo” se ejecuta el proceso en un contexto específico, considerando la secuencia de acciones, los tiempos asignados, las condiciones de aplicación y los criterios previamente establecidos. Una adecuada planificación procedimental resulta fundamental para asegurar la validez, la confiabilidad y la coherencia pedagógica del proceso evaluativo. En esta línea, López-Pastor y Pérez-Pueyo (2017) proponen una serie de “principios de procedimiento” que orientan el diseño e implementación de la evaluación formativa y compartida, destacando la importancia de criterios como la adecuación al currículo, la relevancia de lo evaluado, la veracidad de las evidencias recogidas, y la coherencia entre objetivos, actividades y evaluación. En este sentido, Cruzado (2022) subraya que los procedimientos permiten organizar de forma lógica la aplicación de instrumentos y la recolección de evidencias, especialmente en procesos que priorizan la evaluación formativa. Así, por ejemplo, un procedimiento basado en la técnica del estudio de caso podría incluir la lectura individual del caso, análisis grupal, socialización en plenario y finalmente la entrega de una reflexión escrita.

- *Instrumento*: El instrumento constituye la herramienta concreta mediante la cual se organiza, registra y sistematiza la información obtenida a través de una técnica específica, y es esencial para el análisis objetivo del aprendizaje. Puede adoptar múltiples formas, tales como rúbricas analíticas, listas de cotejo, escalas estimativas, diarios de campo o cuestionarios. Su diseño requiere claridad en los criterios e indicadores de logro, así como una alineación con los objetivos de aprendizaje previamente establecidos. López-Pastor (2009) subraya que los instrumentos hacen visible lo aprendido, facilitando tanto la retroalimentación como la toma de decisiones pedagógicas informadas. Asimismo, Conde-Rojas et al. (2022) enfatizan que un instrumento bien diseñado debe ser sensible al contexto,

confiable en su aplicación y útil para orientar procesos de mejora continua, más allá de la mera medición del desempeño.

La relación entre estos cuatro conceptos puede representarse de manera jerárquica: la estrategia como el plan global; las técnicas como los métodos empleados dentro de esa estrategia; los procedimientos como las secuencias de implementación de las técnicas; y los instrumentos como los medios concretos para registrar y valorar la información obtenida.

Esta delimitación conceptual no solo es útil desde un punto de vista teórico, sino que tiene importantes implicancias en la práctica educativa. Una comprensión clara y diferenciada de estos términos permite al profesorado seleccionar, combinar y aplicar recursos evaluativos de forma más pertinente y coherente, promoviendo procesos más rigurosos y transparentes. En investigaciones recientes, como la de Basurto-Mendoza et al. (2021), se ha destacado la necesidad de integrar estos elementos en el marco de una evaluación formativa y auténtica, centrada en el desarrollo de competencias.

Asimismo, investigaciones como las de Angelini (2016) y Bernard y González-Moreno (2020) han evidenciado que la falta de distinción conceptual entre estrategia, técnica, procedimiento e instrumento puede restringir la capacidad del profesorado para planificar y aplicar procesos evaluativos coherentes, transformando la evaluación en una práctica meramente burocrática o técnica, desprovista de sentido formativo. En síntesis, la articulación clara y reflexiva de estos elementos constituye un pilar fundamental para diseñar evaluaciones pedagógicamente significativas, éticamente responsables y contextualizadas. Esta claridad resulta especialmente crítica en la asignatura de EFyS, donde las decisiones evaluativas deben considerar, sinérgicamente, dimensiones cognitivas, motrices, actitudinales y sociales.

Desde esta base conceptual, se vuelve necesario revisar los distintos tipos de evaluación educativa que han emergido históricamente en el campo pedagógico. Su análisis permite comprender no solo los propósitos que persigue la evaluación, sino también las prácticas y criterios que le dan forma en los contextos escolares actuales.

2.1.5. Tipos de evaluación educativa

A partir de las distinciones teóricas previamente desarrolladas, y con el objetivo de sistematizar la diversidad de formas que puede asumir la evaluación en función de distintos criterios, como el momento de aplicación, el agente evaluador, el enfoque epistemológico subyacente o el propósito del juicio emitido, diversos autores han propuesto clasificaciones tipológicas que permiten comprender su complejidad (Newton, 2007; Perrenoud, 2008). Lejos de constituir un mero ejercicio taxonómico, estas tipologías ofrecen marcos interpretativos que posibilitan analizar cómo se configuran las prácticas evaluativas en función de las intenciones pedagógicas, los contextos institucionales y los marcos teóricos que las sustentan (Pérez et al., 2008; Rodríguez-Gómez y Ibarra-Sáiz, 2011).

Tabla 2. Síntesis de clasificación clásica de la evaluación educativa.

Criterio de clasificación	Tipo de evaluación	Características	Sustento teórico
Temporalidad	Diagnóstica	Se aplica al inicio del proceso para identificar conocimientos previos, habilidades y necesidades del estudiantado.	Bloom et al. (1971); Tyler (2009)
	Formativa	Acompaña el proceso de aprendizaje; busca retroalimentar y ajustar la enseñanza en tiempo real.	Black y Wiliam (2003); Sadler (1989); Harlen y James (1997)
	Sumativa	Se aplica al final del proceso para emitir un juicio sobre el logro de objetivos, generalmente con fines certificativos.	Scriven (1967); Harlen (2005)
Agente evaluador	Autoevaluación	El propio protagonista juzga su desempeño, reflexionando sobre sus logros, dificultades y procesos.	Rodríguez-Gómez y Ibarra-Sáiz (2011); López-Pastor (2011)
	Coevaluación	Evaluación compartida entre agentes que no ocupan la misma posición jerárquica, como un docente y su estudiante, en un proceso negociado y horizontal.	Panadero y Jonsson (2013); Gallardo-Fuentes et al. (2022); López-Pastor et al. (2006)
	Evaluación entre pares	Evaluación que se realiza entre personas que ocupan el mismo rol o estatus, como estudiantes entre sí. Favorece la autorregulación, el juicio crítico y el aprendizaje colaborativo.	Boud y Molloy (2013); Hortigüela-Alcalá et al. (2019); Contreras (2018)
	Heteroevaluación	El docente o evaluador externo emite el juicio sobre el desempeño del estudiante.	Newton (2007); Santos-Guerra (2014)
Enfoque teórico	Cuantitativa	Se basa en la medición objetiva, análisis estadístico y generalización de resultados.	Nitko y Brookhart (2013); Cresswell (2003)
	Cualitativa	Busca interpretar significados, experiencias y contextos desde una lógica inductiva y comprensiva.	Denzin y Lincoln (2011); Guba y Lincoln (1994)
	Mixta	Combina estrategias cualitativas y cuantitativas en función del problema investigado.	Creswell y Plano Clark (2018); Greene (2007)

Elaboración propia

En consecuencia, se presenta a continuación la tabla 2, que reúne los principales tipos de evaluación identificados en la literatura especializada, junto con sus características clave y referentes conceptuales. Este recurso tiene como propósito orientar tanto el análisis crítico como el diseño de evaluaciones pedagógicamente coherentes, contextualizadas y éticamente fundamentadas.

La clasificación de los tipos de evaluación en el ámbito educativo ha sido históricamente objeto de debate y revisión, tanto por su relevancia teórica como por sus implicaciones prácticas en la enseñanza y el aprendizaje. Desde una perspectiva clásica, se ha establecido una distinción fundamental entre evaluación diagnóstica, formativa y sumativa, en función del momento en que se aplica y el propósito que persigue (Bloom et al., 1971; Scriven, 1967). Esta tipología, ampliamente difundida, fue posteriormente refinada por autores como Sadler (1989), quien señaló que la distinción esencial entre evaluación formativa y sumativa no radica en el instrumento utilizado ni en su momento de aplicación, sino en el uso que se hace del juicio evaluativo. Desde esta perspectiva, la evaluación formativa se orienta a mejorar los aprendizajes en curso, mientras que la sumativa tiene como finalidad emitir un juicio terminal sobre el desempeño, generalmente con fines de certificación o acreditación (Black y Wiliam, 2003; Harlen y James, 1997). Esta visión funcionalista ha sido complementada por un enfoque más pragmático, que sitúa a la evaluación dentro de un continuo y reconoce que un mismo instrumento puede ser utilizado con fines distintos según el contexto (Taras, 2005; Wiliam, 2004).

Más allá de esta clasificación, otros enfoques han propuesto categorizar la evaluación según el agente que la ejecuta. Así, se distinguen la autoevaluación, la coevaluación y la heteroevaluación, prácticas que no solo implican distintos roles para docentes y estudiantes, sino que también configuran relaciones de poder diferenciadas en el aula (Del Pino-Sepúlveda y Montanarez-Vargas, 2023). La autoevaluación promueve la reflexión y la autorregulación del aprendizaje, otorgando al estudiantado un rol activo y metacognitivo. La coevaluación, por su parte, fomenta el diálogo horizontal y la construcción colectiva del juicio evaluativo entre pares. En tanto, la heteroevaluación mantiene la lógica vertical tradicional, en la que el docente emite un juicio sobre el desempeño del estudiante (Basurto-Mendoza et al., 2021). Este enfoque triangular de la evaluación ha sido valorado como una

vía innovadora para democratizar los procesos evaluativos y reforzar su carácter formativo, especialmente cuando se implementa de manera integrada y coherente con los objetivos de aprendizaje (Rodríguez-Gómez e Ibarra-Sáiz, 2011).

En una revisión crítica de estas categorías, Newton (2007) propone repensar la noción de "tipo de evaluación" distinguiendo entre niveles de análisis: el nivel del juicio (tipo de resultado que se genera), el nivel de decisión (uso que se le da a dicho juicio) y el nivel de impacto (efectos esperados del sistema evaluativo en el sistema educativo). Según el autor, la clasificación dicotómica entre evaluación formativa y sumativa ha inducido a errores conceptuales, al ubicarse en niveles discursivos distintos. Mientras la evaluación formativa corresponde a un uso de la información (nivel decisonal), la evaluación sumativa se refiere al carácter del juicio evaluativo (nivel de juicio). Esta confusión ha sido replicada en múltiples documentos de política educativa, que intentan agrupar bajo la categoría de "evaluación sumativa" a una variedad de propósitos disímiles, desde la selección hasta la rendición de cuentas (Harlen, 2005; Newton, 2007). En respuesta a esta reducción conceptual, se ha propuesto una tipología más fina, que incluye propósitos como la orientación vocacional, la asignación de recursos, la mejora institucional, la acreditación, entre otros, reconociendo que cada uno de estos fines exige diseños evaluativos específicos y criterios de validez diferenciados (Newton, 2010; Pellegrino et al., 2001).

A medida que los entornos educativos se transforman con la irrupción de tecnologías digitales y herramientas de inteligencia artificial, la tipología de la evaluación también se ha visto interpelada. El uso de sistemas inteligentes ha ampliado el repertorio de formatos evaluativos, posibilitando la retroalimentación inmediata, la personalización de tareas, y la evaluación continua y adaptativa (Mao et al., 2024). Esta expansión, sin embargo, conlleva desafíos éticos y pedagógicos. Por un lado, el uso de algoritmos puede consolidar formas de evaluación automatizada que refuercen la lógica sumativa y la vigilancia, alejándose de los principios de equidad y reflexión crítica (Lee, 2023; Popenici y Kerr, 2017). Por otro lado, emergen oportunidades para repensar la evaluación como un proceso más dialogante, distribuido y centrado en el aprendizaje, siempre que se garantice una apropiación crítica por parte del profesorado y del estudiantado (Cardon et al., 2023). En este contexto, más que abandonar las categorías clásicas, se hace necesario complejizarlas y contextualizarlas,

articulando tipologías funcionales con una reflexión ética y política sobre los fines de la educación y la naturaleza del juicio evaluativo (Del Pino-Sepúlveda y Montanarez-Vargas, 2023; Rodríguez-Gómez y Ibarra-Sáiz, 2015).

En sintonía con estos debates, durante las últimas décadas han emergido enfoques evaluativos que buscan transformar la función de la evaluación desde una lógica exclusivamente certificadora hacia una perspectiva más formativa, participativa y orientada al aprendizaje. Uno de los más relevantes es la evaluación formativa y compartida, entendida como un proceso continuo, negociado y dialógico entre docentes y estudiantes, que promueve la autorregulación del aprendizaje, la retroalimentación constructiva y la corresponsabilidad pedagógica (López-Pastor et al., 2006; Rodríguez-Gómez y Ibarra-Sáiz, 2011). Este modelo enfatiza el papel activo del estudiantado en la toma de decisiones evaluativas, incorporando prácticas como la coevaluación, la autoevaluación y la reflexión metacognitiva. De forma complementaria, la evaluación auténtica ha ganado fuerza como una alternativa que busca establecer vínculos significativos entre el conocimiento escolar y las competencias requeridas en contextos reales. Esta propuesta promueve tareas integradoras, contextualizadas y con sentido para el alumnado, en contraposición a evaluaciones fragmentadas, descontextualizadas y basadas en la reproducción de contenidos (Brown, 2015; García-Gámez, 2024). Asimismo, la evaluación por competencias ha impulsado una mirada integral de los aprendizajes, considerando no solo saberes declarativos, sino también habilidades prácticas, actitudes y capacidades de transferencia a nuevos contextos (Sánchez y Orozco, 2020). Estas aproximaciones comparten una base ética y pedagógica que reconoce a la evaluación como una oportunidad de aprendizaje y como una práctica social situada, con capacidad para empoderar al estudiantado y transformar las culturas escolares desde una lógica más inclusiva, reflexiva y emancipadora.

2.2. *Paradigmas en la evaluación educativa*

En el campo de la investigación educativa suele observarse una cierta confusión conceptual entre los términos paradigma y enfoque. Aunque están relacionados, no son equivalentes. Un paradigma constituye una cosmovisión epistemológica, ontológica y metodológica amplia que orienta la comprensión de los fenómenos educativos, estableciendo supuestos teóricos que guían la producción del conocimiento (Guba y Lincoln, 1994). Por su

parte, un enfoque corresponde a una orientación metodológica o didáctica más específica, que se traduce en decisiones prácticas concretas tales como estrategias, procedimientos o técnicas aplicadas en la evaluación (Carbajosa, 2011). En otras palabras, el paradigma proporciona el marco de sentido desde el cual se interpreta la realidad educativa, mientras que el enfoque traduce ese marco en estrategias prácticas aplicadas en contextos específicos, como lo es la evaluación. Esta distinción es clave para comprender que los distintos enfoques de evaluación, como el formativo y compartido, el auténtico, el basado en competencias o el enfoque multimodal mediado por tecnologías emergentes, se inscriben en marcos paradigmáticos más amplios que configuran concepciones sobre el aprendizaje, el rol del profesorado y los fines de la educación. Como señalan Carbajosa (2011) y López-Pastor y Sicilia-Camacho (2017), los enfoques evaluativos no son neutros ni meramente técnicos, sino que expresan opciones éticas, epistemológicas y pedagógicas sobre cómo se entiende el proceso educativo y para qué se evalúa. Cada uno de estos enfoques responde a una manera particular de concebir la enseñanza y el aprendizaje, orientando no solo los procedimientos evaluativos, sino también los valores pedagógicos que los sustentan.

Desde esta perspectiva es posible identificar, al menos, dos grandes paradigmas clásicos en la evaluación educativa: el paradigma tradicional y el paradigma alternativo. El paradigma tradicional, también denominado técnico o positivista, se basa en una racionalidad instrumental que prioriza la objetividad, la medición estandarizada y la clasificación de los aprendizajes mediante calificaciones cuantitativas (Blázquez, 1990; López, 2006). En este paradigma, el profesorado asume el rol exclusivo de evaluador, el estudiantado es una persona pasiva que responde a pruebas diseñadas externamente, y el objetivo de la evaluación se limita a certificar aprendizajes de forma jerarquizada. La evaluación adquiere así una función principalmente sumativa, de carácter selectivo y controladora, subordinada a criterios de eficacia institucional y reproducción curricular (Perrenoud, 2008).

Frente a ello, el paradigma alternativo se articula desde una racionalidad práctica y dialógica, comprendiendo la evaluación como un proceso formativo, participativo y ético, centrado en el mejoramiento continuo del aprendizaje y en la implicación del propio estudiante en dicho proceso (López-Pastor, 2009; Velázquez y Hernández, 2004). Esta visión reconoce múltiples funciones de la evaluación; diagnóstica, formativa, reguladora y

sumativa, y promueve prácticas como la autoevaluación, la coevaluación y la evaluación compartida (Fernández-Balboa, 2005). Desde esta perspectiva, el conocimiento se construye de forma social e intersubjetiva, y la evaluación se transforma en un acto de reflexión crítica, democratización del aula y construcción de significados compartidos (Díaz, 2005; Prieto, 2015). Como sintetiza Santos-Guerra (1995), las diferencias entre los paradigmas evaluativos no solo se manifiestan en los instrumentos que emplean, sino también en sus fundamentos éticos, políticos y pedagógicos, los cuales revelan concepciones profundamente divergentes sobre el aprendizaje, la enseñanza y la justicia educativa.

En la última década ha comenzado a configurarse un paradigma emergente en la evaluación educativa, estrechamente vinculado al desarrollo y uso de tecnologías digitales avanzadas, en particular aquellas asociadas a la IA, como ChatGPT y otros modelos de lenguaje. Este nuevo escenario no solo amplía las posibilidades técnicas para evaluar, sino que también plantea interrogantes éticos, pedagógicos y epistemológicos inéditos, tensionando los supuestos tanto del paradigma tradicional como del alternativo (Caldeiro, 2024; Zapata-Ros, 2024).

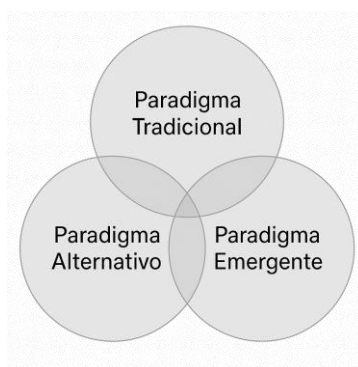
A diferencia del paradigma técnico, centrado en el control externo y la validación objetiva de los aprendizajes, y del paradigma alternativo, que pone énfasis en la participación, la reflexión crítica y la dimensión humana del juicio pedagógico, este paradigma emergente se sitúa en una encrucijada conceptual. La colaboración entre seres humanos y sistemas inteligentes redefine los procesos de evaluación, afectando no solo el qué y cómo se evalúa, sino también quién evalúa y con qué criterios (Zawacki-Richter et al., 2019). En este contexto, la retroalimentación automatizada, la personalización de tareas, la analítica de aprendizaje y la generación de contenidos mediante IA están modificando profundamente la lógica evaluativa tradicional. Estas innovaciones permiten procesos más adaptativos y centrados en el estudiante, ampliando el repertorio de formatos evaluativos disponibles y posibilitando intervenciones más inmediatas y personalizadas (Kasneci et al., 2023).

Herramientas como la retroalimentación automatizada, la analítica de aprendizaje y la generación de contenidos ofrecen oportunidades para diversificar formatos y personalizar procesos, pero también plantean desafíos en torno a la autoría, la equidad, la privacidad y la

agencia docente (Essien et al., 2024; Michel-Villarreal et al., 2023). En este contexto, autores como Zapata-Ros (2024) y Caldeiro (2024) llaman a construir marcos pedagógicos que integren la IA sin desdibujar el juicio docente, resguardando la autenticidad y la intencionalidad de la evaluación.

Por ello, se requiere una alfabetización crítica en IA que permita a estudiantes y docentes comprender el funcionamiento y las implicancias de estas tecnologías. Tal como proponen Cardon et al. (2023), la evaluación no debe ser subordinada a la eficiencia algorítmica, sino fortalecida por una orientación ética y formativa que preserve su sentido pedagógico.

Figura 1.Paradigmas en evaluación educativa.



Elaboración propia

En este contexto, más que hablar de una sustitución de paradigmas, se sugiere un proceso de hibridación que articula elementos de los tres marcos: la necesidad de rigor técnico del paradigma tradicional, la orientación formativa y dialógica del paradigma alternativo y las posibilidades de personalización y análisis que ofrece la inteligencia artificial. Este enfoque híbrido ha sido evidenciado en estudios recientes, como el de Pérez-Pueyo et al. (2017), quienes proponen en el ámbito de la EF una combinación de modelos pedagógicos que favorece prácticas evaluativas más formativas, contextualizadas y centradas en el estudiante. De forma similar, Mantilla-Falcón et al. (2020) documentan en la educación superior ecuatoriana una hibridación de modelos pedagógicos en la práctica docente, que responde a la complejidad y diversidad de contextos educativos contemporáneos. Estos hallazgos respaldan la necesidad de repensar los marcos evaluativos desde una perspectiva ética, situada y crítica, que ponga al servicio del aprendizaje no solo las herramientas tecnológicas, sino también los valores que sustentan una educación humanizadora, reflexiva y equitativa.

Tabla 3. Dimensiones comparativas de los paradigmas de evaluación educativa.

Criterio	Paradigma Tradicional	Paradigma Alternativo	Paradigma Emergente
Propósitos	Certificar, verificar logros, acreditar aprendizajes (Scriven, 1967).	Orientar el aprendizaje, acompañar procesos, favorecer la autorregulación (Black y Wiliam, 2003; López Pastor, 2009).	Adaptación personalizada, monitoreo continuo y retroalimentación automatizada (Kasneci et al., 2023; Michel-Villarreal et al., 2023).
Rol del estudiantado	Pasivo, receptor de información. Su función es responder (Perrenoud, 2008).	Activo, reflexivo y corresponsable del proceso evaluativo (Rodríguez-Gómez y Ibarra-Sáiz, 2011).	Colaborador con sistemas de IA, requiere alfabetización digital crítica (Nguyen et al., 2024; Cardon et al., 2023).
Rol del profesorado	Juez, emisor del juicio evaluativo final. Controla el proceso (Blázquez, 1990).	Facilitador, mediador del aprendizaje. Emite retroalimentación formativa (Santos Guerra, 1995; Díaz, 2005).	Diseñador de experiencias evaluativas con IA, cuidador de la integridad pedagógica (Essien et al., 2024).
Retroalimentación	Escasa, centrada en el resultado (nota numérica) (Harlen y James, 1997).	Permanente, dialógica y personalizada (Sadler, 1989; Allal, 2013).	Instantánea, automatizada y adaptativa; requiere validación humana (Mao et al., 2024; Zawacki-Richter et al., 2019).
Instrumentos utilizados	Pruebas objetivas, de selección múltiple, listas cerradas (Casanova, 1998).	Rúbricas, portafolios, coevaluaciones, diarios reflexivos (López-Pastor et al., 2006; Basurto-Mendoza et al., 2021).	Plataformas de evaluación con IA: retroalimentación automatizada, dashboards de aprendizaje, sistemas de analítica predictiva, entornos adaptativos personalizados (Kasneci et al., 2023; Michel-Villarreal et al., 2023; Crompton y Burke, 2023; Bulut et al., 2024).
Nivel cognitivo	Predomina la reproducción, la memoria y la respuesta única (Bloom et al., 1971).	Comprensión profunda, aplicación, análisis, metacognición (Talanquer, 2015; Perrenoud, 2008).	Evaluación de habilidades complejas, creativas y adaptativas. Riesgo de sesgo automatizado (Kasneci et al., 2023; Zawacki-Richter et al., 2019).

Elaboración propia

En síntesis, los tres paradigmas analizados en la tabla 3; tradicional, alternativo y emergente, configuran marcos interpretativos que orientan de manera diferenciada las finalidades de la evaluación, el rol del profesorado y del estudiantado, los instrumentos utilizados y los niveles cognitivos promovidos. Estas concepciones no solo reflejan distintas visiones sobre el aprendizaje y la función de la escuela, sino que también expresan posturas éticas, epistemológicas y políticas sobre el acto evaluativo.

En el caso del paradigma emergente, su configuración responde a los profundos cambios introducidos por la incorporación de tecnologías digitales avanzadas, particularmente aquellas asociadas a la inteligencia artificial generativa. Diversos autores han

planteado que estas transformaciones no representan únicamente innovaciones instrumentales, sino que están dando lugar a una reformulación conceptual de la evaluación. Por ejemplo, Zapata-Ros (2024) sostiene que la irrupción de herramientas como ChatGPT desafía los modelos pedagógicos existentes y exige una teoría del aprendizaje específica que sustente prácticas evaluativas pertinentes. De manera similar, Caldeiro (2024) plantea que la IAg no solo redefine el rol docente, sino también las condiciones de posibilidad del juicio evaluativo, al introducir nuevas formas de generación de evidencia y retroalimentación automatizada. Ambos autores coinciden en que estamos ante un cambio de paradigma que obliga a repensar no solo los métodos, sino los fundamentos mismos de la evaluación.

No obstante, para comprender cómo estos marcos se materializan en la práctica educativa, resulta necesario profundizar en los propósitos y enfoques de evaluación, entendidos como orientaciones metodológicas más específicas que operacionalizan las ideas paradigmáticas en decisiones concretas de diseño, aplicación e interpretación. Como señala Carbajosa (2011), los enfoques constituyen un plano intermedio entre los supuestos teóricos generales de los paradigmas y las estrategias particulares aplicadas en contextos educativos concretos. En este sentido, el siguiente apartado se centrará en el análisis de diversos propósitos evaluativos, para posteriormente dar pie al desarrollo de los enfoques, considerando que su configuración responde a los fundamentos de los paradigmas previamente descritos, pero también a las condiciones institucionales, contextuales y pedagógicas en las que se insertan.

2.3. Propósitos de la evaluación educativa

La comprensión de los propósitos de la evaluación educativa constituye un eje central para el diseño, implementación y análisis de cualquier sistema evaluativo, ya que de ellos dependen tanto los procedimientos empleados como los usos que se otorgan a los resultados. Tradicionalmente, estos propósitos se han clasificado de forma binaria en términos formativos y sumativos (Black y Wiliam, 2003; Bloom et al., 1971), sin embargo, esta dicotomía ha sido ampliamente cuestionada por su carácter simplificador y por las ambigüedades conceptuales que conlleva (Newton, 2007; Taras, 2005). Desde una perspectiva crítica, Newton (2007) propone distinguir tres niveles de análisis para entender el propósito de la evaluación: el nivel del juicio (tipo de resultado), el nivel de decisión (uso

de los resultados) y el nivel de impacto (efectos deseados del sistema evaluativo). Esta distinción permite superar enfoques reduccionistas y visibilizar una diversidad de finalidades, como el monitoreo del sistema, la orientación pedagógica, la acreditación o la selección (Newton, 2007; Sadler, 1989). En este sentido, Harlen (2005) y Wiliam (2004) enfatizan que lo que otorga carácter formativo a una evaluación no es su formato o momento, sino el uso que se hace de la información recogida. Por ello, resulta más pertinente hablar de finalidades múltiples y situadas, definidas según los actores involucrados, los contextos institucionales y las decisiones pedagógicas implicadas (Harlen y James, 1997; Pellegrino et al., 2001).

2.3.1. Evaluación con propósito de diagnóstico

Uno de los propósitos fundamentales de la evaluación educativa es el diagnóstico de los aprendizajes, entendido como la identificación temprana de fortalezas, debilidades y necesidades del estudiantado. Esta función diagnóstica permite al profesorado obtener una imagen inicial del estado de conocimientos, habilidades y actitudes del alumnado, lo que facilita la toma de decisiones pedagógicas pertinentes y ajustadas al contexto. Tal como establece el Decreto N.º 67 del Ministerio de Educación de Chile, la evaluación debe ofrecer información relevante para retroalimentar oportunamente el proceso de enseñanza y aprendizaje, apoyando decisiones que favorezcan el desarrollo integral del estudiantado (Ministerio de Educación, 2018).

Desde una perspectiva formativa, la evaluación diagnóstica no busca calificar ni certificar, sino comprender el punto de partida del aprendizaje. Esta concepción ha sido respaldada por investigaciones que destacan su valor para orientar una enseñanza diferenciada, inclusiva y centrada en el progreso individual (López Pastor et al., 2006; Rosales, 2014). En estudios recientes sobre prácticas docentes, se ha observado que los y las profesoras reconocen este tipo de evaluación como un recurso clave para ajustar los objetivos de enseñanza a las características particulares de sus estudiantes, promoviendo procesos más equitativos y contextualizados.

Además, el diagnóstico puede revelar factores externos al rendimiento académico que inciden en el proceso formativo, tales como contextos emocionales, niveles de motivación, o

barreras específicas de aprendizaje, elementos esenciales para planificar intervenciones pedagógicas más eficaces. En este sentido, entender la evaluación como un proceso que antecede a la instrucción, y no como una etapa posterior, implica asumir un rol más reflexivo y estratégico del profesorado (Gómez et al., 2006).

2.3.2. Evaluación para rendición de cuentas y mejora institucional

La evaluación educativa cumple, además de sus funciones pedagógicas directas, un papel estratégico en los sistemas escolares contemporáneos al actuar como herramienta para la rendición de cuentas y la mejora institucional. En este ámbito, los resultados de las evaluaciones se utilizan para monitorear el desempeño de los centros educativos, establecer estándares de calidad, orientar decisiones administrativas, asignar recursos o justificar políticas públicas (Newton, 2007; OCDE, 2013). Dicho así, se trata de un propósito ubicado en el nivel de impacto, dado que busca influir más allá del aula, afectando las estructuras del sistema educativo en su conjunto (Newton, 2007).

No obstante, este tipo de evaluación ha sido objeto de debate por su posible instrumentalización y sus efectos adversos. Algunos estudios señalan que, cuando la rendición de cuentas se vincula estrechamente con mecanismos de presión y consecuencias punitivas, se generan efectos contraproducentes como la “enseñanza al test”, la estandarización del currículo, la fragmentación de los aprendizajes y el aumento del estrés docente (Creswell, 2003; Harlen, 2005). En el contexto latinoamericano, Ravela et al. (2016) han advertido que las evaluaciones externas de gran escala no siempre se articulan de forma efectiva con los procesos escolares, y que su impacto positivo depende en gran medida de cómo se utilizan los resultados para generar mejoras en la práctica pedagógica y el desarrollo profesional docente.

Desde una perspectiva crítica, autores como Fuentes-Vásquez et al. (2025) han problematizado el vínculo entre gobernanza educativa y evaluación en Chile, alertando sobre el riesgo de que el discurso de calidad se traduzca en una lógica de control técnico que desatienda las condiciones estructurales de desigualdad. Por ello, se ha insistido en la necesidad de reorientar este tipo de evaluación hacia un enfoque de mejora formativa

institucional, basado en la retroalimentación, la participación del profesorado y el uso reflexivo de la información (Ahumada, 2005; Pérez Pueyo et al., 2008).

En este sentido, para que la evaluación con fines de mejora y rendición de cuentas cumpla efectivamente un rol transformador, debe estar acompañada por políticas que promuevan el fortalecimiento de las capacidades evaluativas de las comunidades escolares, asegurando su apropiación crítica y contextualizada. Lejos de ser solo un instrumento de supervisión, la evaluación puede contribuir a la construcción de culturas organizacionales orientadas al aprendizaje, siempre que se integre con coherencia en los procesos institucionales y pedagógicos.

2.3.3. Evaluación con fines de orientación y apoyo al aprendizaje

La evaluación con fines de orientación y apoyo al aprendizaje constituye uno de los propósitos más relevantes y transformadores dentro del ámbito educativo. Lejos de limitarse a certificar o calificar, este enfoque convierte la evaluación en una herramienta que favorece el aprendizaje, al generar instancias de retroalimentación continua, fomentar la autorregulación y estimular la reflexión metacognitiva del estudiantado (Black y Wiliam, 1998; Fraile et al., 2020). En este marco, la evaluación se convierte en una estrategia pedagógica activa, que no solo informa sobre el desempeño, sino que lo acompaña y lo potencia.

Este enfoque se traduce conceptualmente en lo que se ha denominado *evaluación para el aprendizaje*, una modalidad que sitúa al estudiantado en el centro del proceso evaluativo, promoviendo su participación activa en la definición de criterios, la autoevaluación y la coevaluación (Andrade y Heritage, 2017; López-Pastor et al., 2006). La EFyC, como modalidad consolidada de este enfoque, ha demostrado su eficacia tanto en contextos escolares como universitarios, particularmente en la formación inicial docente (Gallardo-Fuentes et al., 2019; Añazco-Martínez et al., 2024).

López-Pastor et al. (2011) destacan que cuando los y las estudiantes participan activamente en el proceso evaluativo, se genera una comprensión más profunda de los objetivos de aprendizaje, así como un mayor compromiso con sus trayectorias formativas. Esta participación no solo mejora el rendimiento académico, sino que también fortalece

habilidades de pensamiento crítico, colaboración y autogestión. A su vez, Fernández-Garcimartín (2023) ha evidenciado que la EFyC aplicada a los TFG permite el desarrollo de competencias transversales clave como la reflexión, el análisis crítico y la autonomía intelectual.

En el caso de la EF, donde históricamente ha prevalecido una lógica técnica y clasificatoria, la implementación de modelos de evaluación centrados en el proceso y la participación ha permitido ampliar las dimensiones valoradas, incluyendo aspectos actitudinales, sociales y emocionales del aprendizaje (Cañadas et al., 2018; Hortigüela-Alcalá et al., 2019). Este giro implica un cambio de paradigma, en el que el profesorado asume un rol más orientador que sancionador, generando ambientes de aprendizaje más inclusivos y significativos.

Además, investigaciones como la de Barba-Martín et al. (2020) y Fuentes-Diego y Salcines-Talledo (2018) sostienen que la evaluación formativa, al ser implementada de manera continua y reflexiva, contribuye a crear una cultura evaluativa más democrática y colaborativa. Este proceso no está exento de desafíos: su implementación requiere planificación, formación docente y tiempo pedagógico. Sin embargo, sus beneficios en términos de compromiso estudiantil, desarrollo de competencias y mejora del clima de aula son ampliamente documentados.

2.3.4. Nuevos propósitos de la evaluación en la era digital e inteligencia artificial

La transformación digital y el desarrollo acelerado de la IA están configurando un nuevo panorama para la evaluación educativa, replanteando sus propósitos, alcances y modalidades. Tecnologías como la inteligencia artificial generativa, los sistemas de aprendizaje adaptativo, la retroalimentación automatizada y el análisis predictivo del rendimiento abren posibilidades para procesos de evaluación más personalizados, escalables y continuos (Mao et al., 2024; Zawacki-Richter et al., 2019). En este nuevo escenario, la evaluación ya no se limita a certificar aprendizajes pasados, sino que se proyecta como una herramienta para acompañar trayectorias formativas, anticipar necesidades y facilitar decisiones pedagógicas basadas en datos (Chen et al., 2023).

Este enfoque tecnopedagógico, sin embargo, también plantea importantes desafíos éticos y epistemológicos. Por un lado, el uso de algoritmos opacos puede derivar en prácticas de evaluación automatizada descontextualizadas y excluyentes, que reproducen sesgos y refuerzan la lógica sumativa, afectando la equidad y el sentido pedagógico del proceso (Bulut et al., 2024; Popenici y Kerr, 2017; Williamson et al., 2020). Por otro lado, estas mismas tecnologías pueden ser resignificadas desde una perspectiva crítica y educativa, al servicio de una evaluación más inclusiva, formativa y centrada en el aprendizaje. Cardon et al. (2023) proponen, por ejemplo, modelos de evaluación multimodal apoyados en IA que favorecen una retroalimentación continua y situada, basada en la interacción entre docentes, estudiantes y sistemas inteligentes.

Desde este enfoque, la IA no reemplaza al juicio pedagógico, sino que lo complementa, ampliando las formas de evidenciar el aprendizaje y de generar procesos reflexivos con sentido educativo. El desafío, por tanto, no radica únicamente en adoptar nuevas tecnologías, sino en reconfigurar los propósitos de la evaluación en función de principios de justicia educativa, autonomía estudiantil y sentido pedagógico. Como afirman Zawacki-Richter et al. (2019), el verdadero potencial de la IA en evaluación se alcanza cuando las decisiones automatizadas se subordinan a una visión crítica y humanista de la enseñanza.

2.4. *Enfoques de evaluación educativa*

En el marco de los paradigmas que orientan la evaluación educativa, los enfoques constituyen concreciones metodológicas y didácticas que traducen en la práctica los supuestos epistemológicos y pedagógicos dominantes. Mientras un paradigma ofrece una visión amplia sobre el conocimiento, el aprendizaje y la enseñanza (Guba y Lincoln, 1994), los enfoques operan como guías que orientan las decisiones evaluativas tanto en el aula como en las políticas institucionales (Carbajosa, 2011). Desde esta perspectiva, enfoques como el formativo, auténtico, por competencias o multimodal no son entidades autónomas, sino expresiones coherentes con marcos paradigmáticos más amplios, tales como el técnico, el interpretativo o el crítico.

Durante décadas, la evaluación en contextos escolares ha estado marcada por un enfoque técnico centrado en la medición objetiva, la calificación numérica y el control del

rendimiento. Sin embargo, el desarrollo de perspectivas socioconstructivistas y críticas ha impulsado la emergencia de enfoques que resignifican la evaluación como una práctica pedagógica, ética y relacional (López-Pastor et al., 2006; López-Pastor y Sicilia-Camacho, 2017). En este tránsito, se ha ido consolidando una comprensión más compleja y contextualizada de la evaluación, que prioriza el aprendizaje, la participación activa del estudiantado y el juicio profesional docente.

En este proceso de transformación, uno de los enfoques que ha adquirido mayor protagonismo en la literatura especializada y en experiencias concretas de innovación educativa es la evaluación formativa y compartida (EFyC). Este modelo se ha posicionado como una de las contribuciones más relevantes de las últimas décadas en el ámbito de la evaluación educativa, especialmente por su potencial para mejorar el compromiso estudiantil, la autorregulación y la equidad en los procesos de enseñanza-aprendizaje (Bores-García et al., 2021; Fernández Garcimartín et al., 2023; Gallardo-Fuentes et al., 2019). La EFyC no solo implica una mejora técnica en los instrumentos utilizados, sino una transformación profunda del sentido de la evaluación, entendida como oportunidad para aprender y construir conocimiento de forma colaborativa.

Paralelamente, el avance de tecnologías emergentes como la inteligencia artificial generativa ha abierto nuevas posibilidades metodológicas, tales como la retroalimentación automatizada o la personalización de tareas (Kasneci et al., 2023). Sin embargo, estas innovaciones deben ser leídas críticamente y en función de su coherencia con los fines formativos de la evaluación. En este escenario híbrido, donde conviven propuestas tradicionales, innovaciones pedagógicas consolidadas y herramientas tecnológicas emergentes, la evaluación formativa y compartida ofrece un marco robusto y éticamente orientado para guiar las transformaciones actuales.

A continuación, se presenta con mayor profundidad este enfoque, sus fundamentos y su relevancia para una práctica evaluativa más justa, participativa y pedagógicamente significativa.

2.4.1. Evaluación formativa y compartida

La evaluación formativa y compartida constituye uno de los enfoques más transformadores y coherentes con una educación centrada en el aprendizaje y en el desarrollo integral del estudiantado. Se fundamenta en el uso sistemático de la retroalimentación, la implicación activa del alumnado en los procesos evaluativos, la autorregulación y la toma de decisiones pedagógicas orientadas a la mejora continua. Tal como señalan López-Pastor y Pérez-Pueyo (2017), esta modalidad no solo modifica el cómo se evalúa, sino que exige repensar las finalidades de la evaluación, vinculándola con la mejora de la enseñanza, la autonomía del estudiantado y la construcción compartida del conocimiento.

En contraposición con la lógica unidireccional del enfoque tradicional, este modelo se sustenta en una concepción dialógica, ética y participativa de la evaluación (Hortigüela et al., 2019; López-Pastor, 2011). Desde esta perspectiva, evaluar no significa calificar, sino generar información valiosa para la toma de decisiones pedagógicas, favorecer la reflexión metacognitiva y mejorar el proceso de enseñanza-aprendizaje de manera continua y situada. El estudiantado, en este enfoque, se reconoce como sujeto corresponsable de su propio proceso formativo, capaz de participar activamente en la construcción de los criterios, en la observación de su desempeño y en el diálogo sobre su progreso (Gallardo-Fuentes et al., 2019; Suazo, 2023;).

Tal como lo expone López-Pastor (2017), la EFyC se articula a través de cinco líneas estratégicas complementarias: (1) cuaderno del profesor con evidencias sistemáticas, (2) actividades de aprendizaje significativas, (3) producciones del alumnado, (4) procesos de autoevaluación, coevaluación y evaluación compartida, y (5) ciclos de investigación-acción para el perfeccionamiento progresivo del sistema evaluativo. Esta estructura metodológica permite una aplicación versátil y contextualizada en distintos niveles educativos, siempre orientada por criterios de calidad educativa como la adecuación, la relevancia, la veracidad, la integración, la viabilidad y la ética.

La implicación del alumnado no se reduce a ejercicios puntuales de coevaluación, sino que abarca un conjunto amplio de prácticas como la autocalificación, la calificación dialogada, el uso de rúbricas construidas colaborativamente y el portafolio reflexivo,

constituyéndose en una verdadera cultura evaluativa democratizadora (López-Pastor et al., 2006; López-Pastor, 2009). Esta participación activa, además de mejorar el rendimiento académico, fortalece competencias transversales como el pensamiento crítico, la autorregulación, la responsabilidad y la conciencia de aprendizaje (Asenjo-Paredes et al., 2024; Barba-Martín et al., 2020).

2.4.1.1. Evaluación formativa y compartida en educación superior

Diversos estudios han demostrado que la implementación de la EFyC en la educación superior aporta beneficios significativos tanto en los aprendizajes como en los aspectos actitudinales y motivacionales del estudiantado. Esta modalidad promueve una mayor implicación activa, mejora la autorregulación y fortalece el sentido del aprendizaje, al desplazar el foco desde la calificación hacia la comprensión profunda y la mejora continua. En particular, se han evidenciado mejoras en la motivación intrínseca, la percepción de justicia evaluativa y la autonomía del estudiante (Fernández-Garcimartín et al., 2023; López-Pastor y Sicilia-Camacho, 2017; Molina et al., 2020).

En el ámbito universitario, y especialmente en la formación inicial docente, la EFyC ha demostrado ser clave para el desarrollo de competencias profesionales complejas. Permite al estudiantado participar en la elaboración de rúbricas, la gestión de portafolios reflexivos, procesos de coevaluación, autocalificación y calificación dialogada, transformando la evaluación en un espacio de aprendizaje activo y ético (Gallardo-Fuentes et al., 2017; Herrero-González et al., 2023; Suazo González, 2023).

López-Pastor y Pérez-Pueyo (2017) destacan que es prácticamente imposible formar profesionales comprometidos con prácticas evaluativas formativas si no han tenido experiencias concretas de EFyC durante su formación inicial. En este sentido, la participación del estudiantado en actividades de evaluación entre pares, autoevaluación estructurada y análisis conjunto de evidencias no solo impacta en su rendimiento, sino también en su capacidad de transferir estos modelos a su futura docencia.

Adicionalmente, investigaciones recientes han revelado que la EFyC contribuye a mejorar el rendimiento académico, disminuir el abandono escolar y fomentar una cultura institucional centrada en el aprendizaje, siempre que su implementación vaya acompañada

de formación docente y espacios para la innovación evaluativa (Barba-Martín et al., 2020; Hamodi y López, 2012; Herranz y López-Pastor, 2014; Martos-García y García-Puchades, 2023).

2.4.1.2. Evaluación formativa y compartida en otros niveles educativos

La evaluación formativa y compartida no se limita al ámbito universitario; su implementación ha sido ampliamente documentada en distintos niveles del sistema educativo, evidenciando beneficios significativos tanto para el aprendizaje del estudiantado como para el desarrollo profesional docente (Hamodi et al., 2015; Herrero-González et al., 2024; López-Pastor y Sicilia-Camacho, 2017). En todos estos contextos, la participación activa del alumnado en el proceso evaluativo, combinada con retroalimentación constante, no solo potencia los aprendizajes, sino que también fortalece la autorregulación, la autonomía y un clima de aula más horizontal y dialogante.

En educación infantil, este enfoque ha sido aplicado como una herramienta de acompañamiento respetuosa con los ritmos individuales, favoreciendo la autoexpresión, el reconocimiento de emociones y la reflexión incipiente sobre el propio aprendizaje. Como señalan López Pastor y García Herranz (2017), esta evaluación no solo permite alcanzar los objetivos curriculares, sino que también “despierta en nuestros alumnos actitudes tremendamente valiosas como son la autonomía, dedicación y compromiso hacia la tarea, además de superación personal y autoestima” (p. 124). Asimismo, representa una oportunidad para el crecimiento docente, al entender que “la evaluación [es] un instrumento al servicio del maestro para generar conocimientos y aprendizaje” (López Pastor y García Herranz, 2017, p. 124). A través de recursos adaptados como caritas, dibujos, murales o portafolios colectivos, se promueve una evaluación comprensiva, afectiva y coherente con los principios pedagógicos de la etapa.

En educación secundaria, distintas experiencias han incorporado con éxito prácticas como la autoevaluación, la coevaluación, la calificación dialogada y el uso compartido de rúbricas, estrategias que han demostrado mejorar la motivación, fomentar la responsabilidad del alumnado sobre sus propios procesos y transformar el enfoque sancionador tradicional hacia uno centrado en la mejora continua (Bendezú Monge, 2025; López-Pastor et al., 2006;

Fuentes-Diego y Salcines-Talledo, 2018). Estos dispositivos, además de facilitar una retroalimentación efectiva, refuerzan la implicación del estudiantado en la toma de decisiones evaluativas, favoreciendo su compromiso con el aprendizaje.

En la educación técnico-profesional, la evaluación formativa y compartida ha permitido vincular los procesos evaluativos con la resolución de problemas auténticos, propios del mundo laboral. El desarrollo de proyectos interdisciplinarios, rúbricas colaborativas, observaciones entre pares y portafolios de evidencias ha favorecido una evaluación integral, capaz de valorar no solo el dominio técnico, sino también competencias transversales como el trabajo en equipo, la adaptabilidad y la reflexión crítica (Buitrago et al., 2018; Añazco Martínez et al., 2024).

Más allá de su aplicabilidad operativa, este enfoque ha sido ampliamente reconocido por su capacidad de democratizar las relaciones educativas. La evaluación formativa y compartida promueve vínculos horizontales entre docentes y estudiantes, fomenta procesos éticos e inclusivos y contribuye a una cultura escolar centrada en el aprendizaje y la justicia educativa (Barba-Martín et al., 2020; López-Pastor y Pérez-Pueyo, 2024; Herrero-González et al., 2024). De esta forma, se configura no solo como una estrategia pedagógica, sino también como una práctica transformadora con profundo sentido político y social.

Aunque su implementación requiere tiempo, formación continua y condiciones institucionales favorables, diversos estudios coinciden en que representa una vía eficaz para avanzar hacia una educación más reflexiva, inclusiva y equitativa (Anijovich, 2017; Asenjo-Paredes et al., 2024; Sevil-Serrano et al., 2020). Su desarrollo sostenido puede contribuir a consolidar culturas evaluativas comprometidas con el aprendizaje profundo, la participación activa y el reconocimiento del estudiantado como agente central del proceso formativo.

2.4.1.3. Evaluación formativa y compartida en el contexto chileno

En el contexto chileno, la evaluación formativa y compartida ha comenzado a consolidarse como una práctica pedagógica relevante tanto en el sistema escolar como en la formación inicial docente. En el ámbito universitario, Gallardo-Fuentes et al. (2017) identificaron indicios de apropiación progresiva de este enfoque en una universidad pública, señalando una alta coherencia entre los sistemas evaluativos y las planificaciones docentes,

aunque también una escasa participación estudiantil en la toma de decisiones. En un estudio posterior, Gallardo-Fuentes y Carter-Thuillier (2016) analizaron su implementación en el prácticum de la carrera de EF, constatando que el estudiantado valoró positivamente esta modalidad por su carácter justo, dialógico y centrado en el proceso, aunque persistieron ciertas tensiones entre las percepciones de los estudiantes y los tutores sobre el apoyo en aula.

Una experiencia particularmente significativa es la reportada por Gallardo-Fuentes et al. (2025), quienes compararon dos cohortes de estudiantes en práctica y formadores de tres universidades chilenas. Sus hallazgos revelan una progresiva incorporación de instrumentos como portafolios y rúbricas formativas, especialmente en cohortes recientes, aunque la toma de decisiones evaluativas continúa centrada en el profesorado. Esta tensión entre intencionalidad pedagógica y prácticas heredadas también fue observada en el estudio de Retamal et al. (2023), quienes documentaron una experiencia didáctica en la Universidad de Antofagasta basada en evaluación formativa y compartida. Tanto docentes como estudiantes valoraron la retroalimentación como factor clave en el compromiso con el aprendizaje, aunque se reportaron dificultades en la distribución del liderazgo evaluativo dentro de los grupos.

Complementariamente, Muñoz-Sepúlveda y Valencia-Peris (2022) analizaron una experiencia de simulación pedagógica en la carrera de Pedagogía en EF, evidenciando que el uso de rúbricas con criterios compartidos fortaleció la percepción de justicia evaluativa y la autonomía estudiantil.

En el sistema escolar, los avances también son significativos. Añazco et al. (2024), en una revisión sistemática centrada en América Latina, identificaron múltiples experiencias latinas de EFyC en EF, destacando su efectividad para enseñar habilidades motrices, promover actitudes cooperativas y fomentar la reflexión crítica. En una línea similar, Gallardo-Fuentes et al. (2025) llevaron a cabo un estudio cuasi-experimental con docentes de enseñanza media, constatando que la formación en evaluación formativa y compartida fortaleció su disposición para aplicar prácticas más inclusivas, centradas en el aprendizaje y no en la mera calificación.

Finalmente, Asenjo-Paredes et al. (2024) analizaron las concepciones y prácticas evaluativas en la formación inicial del profesorado de EF, concluyendo que, si bien persisten tensiones con enfoques tradicionales, existe una tendencia creciente hacia modelos más participativos, éticos y situados. Así, la evaluación formativa y compartida en el contexto chileno emerge no solo como una estrategia metodológica eficaz, sino también como una vía para democratizar la enseñanza, promover el juicio crítico y contribuir a una cultura educativa más justa y transformadora.

2.4.2. Evaluación auténtica

Gulikers et al. (2004) definen la evaluación auténtica como un enfoque centrado en tareas que se asemejan a las demandas del mundo real, lo que implica una filosofía educativa subyacente. Este modelo se orienta a la aplicación integrada de conocimientos, habilidades y actitudes, favoreciendo la comprensión significativa y la transferencia del aprendizaje (Sabirón y Arraiz, 2016; Vallejo y Molina, 2014). A diferencia de las pruebas estandarizadas, la evaluación auténtica enfatiza la elaboración de productos complejos, la resolución de problemas reales y el desarrollo de competencias que trascienden los contenidos disciplinares tradicionales.

Las evidencias empíricas coinciden en que este enfoque promueve el desarrollo de habilidades cognitivas de orden superior, tales como la metacognición, la toma de decisiones y el pensamiento crítico, así como una mayor autonomía del estudiantado. Por ejemplo, en el ámbito universitario, se ha empleado con éxito para potenciar el aprendizaje profundo y la reflexión crítica sobre la propia práctica, especialmente en carreras como Educación, Enfermería o Psicología, donde el juicio profesional y la capacidad de actuación situada son elementos fundamentales (Barrientos-Hernán et al., 2020; Trejo, 2024).

La implementación de este enfoque se ha sustentado en estrategias como el uso de portafolios reflexivos, estudios de caso, proyectos interdisciplinarios, tareas colaborativas y simulaciones profesionales, todas ellas orientadas a evidenciar no solo el producto final, sino también los procesos de aprendizaje desarrollados. Asimismo, se ha valorado su potencial para vincular los aprendizajes con desafíos reales del entorno sociolaboral, dando sentido a

la formación académica desde una perspectiva situada y funcional (Sánchez y Orozco, 2020; Trejo González, 2024).

Se han documentado experiencias relevantes que incorporan tareas como la planificación de sesiones didácticas, la enseñanza entre pares, la elaboración de video-tutoriales o la creación de materiales evaluativos por parte del estudiantado, lo que permite evidenciar procesos de aprendizaje complejos y significativos (Fuentes y López-Pastor, 2017; George y González-Moreno, 2020). Estas estrategias fomentan la autorregulación, la reflexión crítica y la apropiación del proceso evaluativo como parte inherente al aprendizaje en EF.

Además, la evaluación auténtica se enriquece mediante el uso de metodologías como la heteroevaluación, la coevaluación y la autoevaluación, las cuales favorecen una retroalimentación horizontal y continua. Lorente-Catalán (2017) subraya que este enfoque no se limita a valorar la habilidad física, sino que incluye competencias como la empatía, la comunicación y la toma de decisiones compartidas, fundamentales para el desarrollo integral del alumnado.

No obstante, su implementación efectiva exige un cambio profundo en la cultura evaluativa de los centros escolares y en las prácticas del profesorado. Como señalan Medina y Díaz (2022), no basta con introducir nuevas metodologías si no se transforman las creencias pedagógicas que sostienen la evaluación como control y castigo. Es necesario diseñar entornos institucionales que favorezcan la experimentación, el trabajo colaborativo y el desarrollo profesional docente en torno a enfoques alternativos de evaluación. En este sentido, la literatura especializada coincide en que la evaluación auténtica requiere condiciones estructurales e institucionales propicias, tiempo para su planificación, criterios consensuados, instrumentos flexibles y una actitud ética y dialógica por parte del profesorado. A pesar de estas exigencias, sus beneficios han sido ampliamente reconocidos, ya que promueve una educación más pertinente, justa y comprometida con la realidad del estudiantado (Barrientos et al., 2023; Bernard y González-Moreno, 2020).

2.4.3. Evaluación por competencias

En estrecha relación con los enfoques formativo y auténtico, la evaluación por competencias ha emergido como un marco integrador que articula conocimientos,

habilidades y actitudes en torno a la resolución de problemas en contextos específicos. Este enfoque desplaza el énfasis desde la mera adquisición de contenidos hacia la capacidad del estudiantado para actuar de forma pertinente, autónoma y reflexiva ante situaciones complejas y cambiantes (García y García, 2022). A diferencia de los modelos centrados exclusivamente en la medición de resultados finales, considera tanto los procesos como los recursos movilizados y las evidencias situadas que dan cuenta del desarrollo personal y profesional (Sánchez y Orozco, 2020).

En la educación superior, la evaluación por competencias se ha promovido en el marco del tránsito hacia un modelo centrado en el estudiante, especialmente influido por el Espacio Europeo de Educación Superior y los procesos de transformación curricular en América Latina. En este contexto, se han desarrollado estrategias e instrumentos que permiten valorar competencias transversales, como el pensamiento crítico, la comunicación efectiva y el trabajo colaborativo, junto con competencias específicas de cada área disciplinar (Fernández-Ferrer y Forés, 2016).

Uno de los desarrollos metodológicos más relevantes ha sido la incorporación del *assessment center*, una estrategia que recrea entornos simulados donde se evalúa el desempeño en tareas contextualizadas a través de rúbricas y observación sistemática. Esta herramienta ha demostrado ser útil para valorar competencias complejas, integrando dimensiones actitudinales y procedimentales del aprendizaje (Estupiñán et al., 2023).

Desde una perspectiva curricular, se han propuesto modelos de trazabilidad de competencias que permiten su seguimiento longitudinal a lo largo del plan de estudios, mediante mapas que integran aprendizajes en sentido vertical (entre niveles) y horizontal (entre asignaturas). Estos modelos, además de favorecer una planificación coherente, posibilitan una evaluación alineada con el perfil de egreso y una retroalimentación pertinente para la toma de decisiones académicas (Sánchez et al., 2018).

Diversas experiencias en formación inicial docente han incorporado estrategias como el uso de portafolios reflexivos, rúbricas compartidas y observación de clases, orientadas a valorar el desempeño en contextos reales y fomentar la reflexión crítica sobre la práctica pedagógica. Este tipo de instrumentos no solo permiten evidenciar el logro de competencias

específicas, sino que también fortalecen la toma de decisiones autónoma y el pensamiento profesional docente (Cañadas et al., 2019; Gallardo et al., 2022).

Sin embargo, la implementación de este enfoque no está exenta de desafíos. Entre los principales obstáculos se encuentran la resistencia cultural a abandonar esquemas tradicionales centrados en la calificación, la sobrecarga docente, la escasa formación específica en evaluación por competencias y la dificultad para operacionalizar ciertos desempeños sin caer en reduccionismos tecnocráticos (Almerich et al., 2018; Zabalza y Lodeiro, 2019). Estas tensiones reflejan la necesidad de repensar no solo los instrumentos evaluativos, sino también las condiciones institucionales y formativas que los sustentan.

Aun así, la literatura especializada destaca que una implementación contextualizada, gradual y pedagógicamente fundamentada puede contribuir a una educación más significativa y orientada al desarrollo integral del estudiantado (Fernández-Ferrer y Forés, 2016; Tejada y Ruiz, 2015). El diseño de sistemas de evaluación articulados con el perfil de egreso, la formación continua del profesorado y la cultura colaborativa en los equipos docentes aparecen como factores clave para sostener procesos evaluativos centrados en competencias reales y transferibles (Espejo et al., 2021; López-Pastor et al., 2019; Muñoz-Sepúlveda et al., 2024). Estas condiciones favorecen una evaluación coherente con los principios del aprendizaje significativo, la autorregulación y la mejora continua, integrando lo formativo con lo institucional de forma crítica y transformadora.

2.4.4. Evaluación digital, inteligente y multimodal

En el escenario educativo contemporáneo, la incorporación de tecnologías digitales y, más recientemente, de IA, ha impulsado un nuevo enfoque en los procesos evaluativos. Este enfoque, que podría denominarse evaluación digital, inteligente y multimodal, se caracteriza por la integración de herramientas tecnológicas avanzadas, sistemas automatizados de análisis y múltiples formatos de representación del conocimiento. Estos desarrollos han ampliado las posibilidades para valorar aprendizajes complejos, permitiendo implementar evaluaciones adaptativas, generar retroalimentación automatizada, monitorear trayectorias de aprendizaje y personalizar el acompañamiento pedagógico (Zawacki-Richter et al., 2019; Mao et al., 2024).

La evaluación digital se apoya en plataformas inteligentes, *dashboards* visuales, y sistemas de analítica del aprendizaje que ofrecen a los docentes datos en tiempo real sobre el progreso del estudiantado. Estos entornos permiten identificar patrones, diagnosticar dificultades, proponer actividades diferenciadas y tomar decisiones pedagógicas con mayor precisión. Asimismo, los sistemas de IA generativa han comenzado a ser utilizados para emitir comentarios automáticos, corregir tareas escritas, sugerir recursos o generar actividades de refuerzo, con una capacidad de escalabilidad inédita en entornos tradicionales.

Sin embargo, el uso de estas tecnologías no está exento de desafíos. Varios autores advierten sobre los riesgos éticos, pedagógicos y epistemológicos asociados al uso acrítico de la inteligencia artificial en contextos educativos. Entre ellos se incluyen la deshumanización del juicio docente, la opacidad de los algoritmos, la reproducción de sesgos estructurales, la vigilancia excesiva y el reemplazo de criterios pedagógicos por lógicas técnicas (Essien et al., 2024; Kasneci et al., 2023). En lugar de potenciar la equidad, algunos sistemas pueden reforzar las desigualdades existentes si no se diseñan e implementan con sentido pedagógico.

Ante este panorama, se vuelve fundamental avanzar hacia una alfabetización digital crítica del profesorado, que permita comprender el funcionamiento de los algoritmos, cuestionar sus implicancias y apropiarse de la tecnología desde un posicionamiento ético, situado y formativo (Cardon et al., 2023; Michel-Villarreal et al., 2023). La formación docente en este ámbito debe ir más allá del uso técnico de plataformas, promoviendo una reflexión sobre los fines de la evaluación, el sentido de los datos producidos y el lugar de la interacción humana en los procesos de retroalimentación y acompañamiento.

En paralelo, la evaluación multimodal ha ganado protagonismo como una estrategia que reconoce la diversidad de formas en que el estudiantado puede construir y expresar conocimiento. Este enfoque propone combinar evidencias escritas, orales, visuales, performativas, digitales y audiovisuales, adaptándose a distintos estilos de aprendizaje, lenguajes y contextos socioculturales. En lugar de privilegiar un único formato, se valora la pluralidad expresiva y se promueve una evaluación más inclusiva, situada y significativa (Ossa y Willatt, 2023).

La multimodalidad no implica solo diversidad de formatos, sino también una concepción ampliada del conocimiento, donde las habilidades comunicativas, la creatividad, el pensamiento visual, la argumentación y la expresión corporal tienen un lugar legítimo en la construcción de evidencias. Esta perspectiva resulta especialmente valiosa en escenarios de inclusión, atención a la diversidad y justicia epistémica, pues permite adaptar la evaluación a las trayectorias, intereses y capacidades del estudiantado.

Ambos enfoques, el digital-inteligente y el multimodal, reflejan un cambio más amplio en las prácticas evaluativas contemporáneas: el tránsito desde una evaluación centrada en la clasificación y el control, hacia modelos más éticos, flexibles y orientados al aprendizaje significativo. Esta transformación implica superar obstáculos institucionales como la sobrecarga docente, la lógica de rendición de cuentas estandarizada, la falta de formación específica y las resistencias culturales al cambio, promoviendo una cultura evaluativa que integre tecnología, pedagogía y justicia educativa (Lleixà et al., 2020; López-Pastor y Pérez-Pueyo, 2024).

2.4.5. Retos y desafíos en la implementación de enfoques

La implementación de enfoques alternativos de evaluación, como la evaluación formativa, auténtica, por competencias o multimodal, conlleva profundas transformaciones en la cultura evaluativa de las instituciones educativas. Si bien sus ventajas pedagógicas están ampliamente documentadas, su aplicación práctica enfrenta importantes obstáculos de orden estructural, cultural y profesional. La literatura especializada destaca que para lograr una implementación sostenida se requieren condiciones institucionales, metodológicas y actitudinales que, en muchos casos, aún no están presentes en los sistemas educativos tradicionales (Angelini, 2016; Brown, 2015).

Uno de los principales desafíos es el cambio de mentalidad docente. Muchos profesores y profesoras se mantienen anclados en enfoques evaluativos tradicionales, centrados en la calificación sumativa y en pruebas objetivas, lo que dificulta la transición hacia modelos más dialógicos y centrados en el proceso. Avanzar hacia una evaluación auténtica o formativa requiere una comprensión profunda del rol de la evaluación como motor de aprendizaje, lo que implica procesos de formación inicial y continua que aborden tanto la dimensión técnica

como la ética y pedagógica (Oltra, 2017; Ossa y Willatt, 2023). En esta línea, diversos estudios coinciden en que la formación docente es una condición indispensable para el cambio evaluativo. La incorporación de estrategias como la autoevaluación, la coevaluación o la retroalimentación efectiva exige competencias pedagógicas específicas, seguridad profesional y referentes de buenas prácticas. La ausencia de formación puede derivar en implementaciones superficiales, malentendidos conceptuales y, en consecuencia, frustración o rechazo tanto por parte del profesorado como del estudiantado (Angelini, 2016; Trejo, 2024).

Otro conjunto de obstáculos está relacionado con los tiempos, espacios y recursos disponibles en las instituciones. La planificación de tareas auténticas, la entrega de retroalimentación personalizada y el seguimiento continuo requieren una inversión considerable de tiempo y espacios, a menudo incompatibles con estructuras escolares marcadas por altas cargas horarias, ratios excesivas o condiciones contractuales precarias. Superar estas barreras exige políticas institucionales que reconozcan a la evaluación como parte sustantiva del proceso formativo y aseguren las condiciones para su desarrollo sostenido (Brown, 2015; Medina y Díaz, 2022).

La coherencia curricular también emerge como un factor clave. La evaluación formativa, auténtica o por competencias solo puede desplegar su potencial si está alineada con los objetivos de aprendizaje, las metodologías de enseñanza y los contenidos del currículo. En ausencia de dicha alineación, las innovaciones evaluativas se tornan frágiles y efímeras, sin capacidad real de transformación. Esto requiere una planificación curricular integral y el compromiso colectivo de los equipos docentes (López-Pastor et al., 2008; Sánchez y Orozco, 2020). Asimismo, la participación del estudiantado no puede ser entendida como un requisito superficial, sino como un componente fundamental del proceso evaluativo. La coevaluación, la retroalimentación entre pares o la autoevaluación requieren de una alfabetización evaluativa que enseñe al alumnado a comprender criterios, usar rúbricas, argumentar decisiones y monitorear su propio progreso. Sin estos aprendizajes, su participación puede volverse simbólica o forzada, perdiendo su sentido formativo (García y López-Pastor, 2015; Suazo, 2023).

Un desafío adicional es la cultura institucional dominante, generalmente orientada a indicadores cuantitativos, rankings y resultados estandarizados como parámetros exclusivos de calidad. Este enfoque colisiona con los principios de los enfoques alternativos, que valoran trayectorias diversas, juicios pedagógicos contextualizados y formas cualitativas de evidencia. Transformar esta lógica requiere legitimar prácticas evaluativas situadas, ampliar las nociones de evidencia y construir una cultura institucional comprometida con la mejora, la inclusión y la equidad (Oltra, 2017; Vallejo y Molina, 2014).

En la actualidad, el uso de tecnologías digitales y herramientas basadas en IA constituye una realidad creciente en los entornos educativos. Lejos de ser ignoradas o descartadas, estas herramientas deben ser comprendidas y orientadas desde marcos pedagógicos críticos, éticos y formativos. Aplicaciones como portafolios electrónicos, plataformas de retroalimentación asincrónica o sistemas de coevaluación automatizada pueden enriquecer la evaluación formativa y compartida, siempre que su implementación responda a propósitos claros, cuente con formación docente adecuada y se desarrolle en condiciones técnicas mínimas (Fuentes y López-Pastor, 2017; Trejo, 2024).

En este sentido, avanzar hacia modelos evaluativos más justos y participativos no depende únicamente de la incorporación tecnológica, sino de un cambio sistémico más profundo. Este implica transformar la cultura profesional docente, promover políticas institucionales sostenibles, reorganizar los tiempos escolares, fortalecer la formación continua y el protagonismo estudiantil, e integrar la tecnología de manera reflexiva como parte de un proceso educativo integral y situado.

2.5. La evaluación en contextos locales

2.5.1. Evaluación educativa en América Latina

Durante las últimas décadas, América Latina ha vivido una profunda reconfiguración de sus sistemas de evaluación educativa, en el marco de un proceso más amplio de reformas orientadas por principios de calidad, eficiencia y estandarización. Estas transformaciones han estado marcadas por la influencia directa de organismos internacionales como la OCDE, la UNESCO y el Banco Mundial, que han promovido un enfoque tecnocrático de la evaluación, orientado a la producción de indicadores comparables, al fortalecimiento de capacidades

institucionales y al monitoreo de resultados de aprendizaje como base para la toma de decisiones educativas (Ferrer, 2006; Murillo y Román, 2010; Zacarias, 2018). Bajo esta lógica, se han implementado reformas estructurales en distintos países de la región, dirigidas a estandarizar pruebas, generar rankings escolares y condicionar políticas educativas a los resultados obtenidos, reforzando una visión gerencial y “rendición de cuentas” de la evaluación.

En este contexto, la evaluación de aprendizajes ha adquirido una centralidad inusitada a través de mecanismos como PISA, el ERCE y diversas evaluaciones nacionales estandarizadas. Estas pruebas, si bien han permitido visibilizar importantes brechas de aprendizaje entre países y al interior de ellos, también han sido objeto de críticas por parte de la comunidad académica y docente. Por ejemplo, los resultados de PISA 2018 en matemáticas mostraron que, aunque Chile y Uruguay fueron los países latinoamericanos con mejor desempeño, ambos se ubicaron significativamente por debajo del promedio OCDE. En contraste, países como Panamá y República Dominicana no alcanzaron siquiera los niveles mínimos de desempeño establecidos (Díaz-Pinzón, 2021). Más allá de los puntajes, estos resultados han desatado debates sobre la pertinencia pedagógica de dichas mediciones, especialmente cuando son utilizadas como único criterio para juzgar la calidad del sistema educativo. Críticos de esta visión argumentan que se corre el riesgo de reducir la complejidad del proceso educativo a cifras estandarizadas, sin considerar los contextos socioculturales, las desigualdades históricas ni la diversidad de trayectorias del estudiantado (Ravela et al., 2016).

Desde una perspectiva crítica, diversos autores han advertido sobre el uso político, instrumental y muchas veces reduccionista de estas evaluaciones estandarizadas. En lugar de constituirse en herramientas para mejorar los procesos de enseñanza y aprendizaje, los sistemas de alta consecuencia tienden a instalarse como dispositivos de control externo que subordinan las decisiones pedagógicas a criterios tecnocráticos. Este tipo de evaluación refuerza una lógica de vigilancia más que de desarrollo profesional docente. Escudero (2010) describe este fenómeno como una forma de panoptismo evaluativo, donde la función de la evaluación se distorsiona hacia la fiscalización, debilitando la autonomía docente y transformando la práctica evaluativa en una estrategia de control jerárquico. En este

escenario, las evaluaciones pierden su sentido formativo, y se transforman en mecanismos punitivos que generan presión sobre las escuelas, especialmente aquellas que atienden a poblaciones más vulnerables, reforzando así las desigualdades estructurales del sistema.

Respecto al desarrollo de sistemas nacionales de evaluación, Ferrer y Fiszbein (2015) sostienen que, si bien se han logrado avances normativos en varios países latinoamericanos, persisten importantes debilidades estructurales que limitan su impacto real. Entre los principales problemas identificados se encuentran la débil articulación entre los marcos curriculares, los estándares de aprendizaje y los sistemas de formación docente. En muchos casos, los dispositivos evaluativos se desarrollan de forma paralela al resto de las políticas educativas, sin una integración orgánica que permita retroalimentar los procesos pedagógicos. Además, la implementación de estos sistemas suele responder más a exigencias de organismos multilaterales o a coyunturas políticas que a una visión educativa propia, coherente y sostenida en el tiempo. Esto ha generado estructuras de evaluación técnicamente robustas, pero desconectadas de los actores escolares y sus necesidades reales.

En el ámbito específico de la evaluación docente, se ha consolidado una tendencia regional hacia esquemas de evaluación centrados en el control del desempeño, antes que en la mejora continua y el desarrollo profesional. Esta lógica, heredera de modelos de gestión empresarial aplicados al sector público, tiende a operar desde una concepción prescriptiva y castigadora, donde el profesorado es objeto de fiscalización más que sujeto de crecimiento. Como resultado, se han documentado efectos adversos como la desmotivación profesional, el incremento del estrés laboral y la subordinación de la práctica pedagógica a la lógica de los indicadores de cumplimiento (Durán, 2017). En este sentido, Ravela et al. (2016) advierten que la sobredimensión de la rendición de cuentas ha desplazado otras funciones clave de la evaluación, como el acompañamiento formativo, la reflexión crítica y el fortalecimiento de la práctica docente. Así, se configura un modelo de evaluación docente que mide, pero no necesariamente mejora.

Frente a este panorama, distintos autores y redes académicas han llamado a reconceptualizar la evaluación desde marcos más éticos, contextuales y pedagógicos. En lugar de reforzar una cultura de la fiscalización, se propone recuperar el sentido formativo,

inclusivo y situado de los procesos evaluativos. En esta línea, Ramírez-Quintero y Rojas-Gómez (2025) plantean una visión de la evaluación centrada en el desarrollo integral del estudiantado, incorporando metodologías activas, herramientas digitales con fines pedagógicos, y una mirada que reconozca la pluralidad de saberes presentes en las aulas. Para ello, es indispensable transitar desde sistemas prescriptivos hacia otros más abiertos, flexibles y participativos, donde tanto el profesorado como el estudiantado sean agentes activos en el proceso evaluativo.

Ravela et al. (2016), por su parte, insisten en que la evaluación debe recuperar su función pedagógica y formadora, favoreciendo procesos de diálogo, reflexión y toma de decisiones compartida. Esta visión choca con la lógica clasificatoria y castigadora que aún prevalece en muchos sistemas de la región, donde la medición del rendimiento se ha instalado como criterio exclusivo de calidad educativa.

En este mismo marco crítico, se ha cuestionado el carácter reproductor de desigualdades que asumen muchos de los modelos evaluativos vigentes. Zacarias (2018) advierte que el énfasis en la medición, sin considerar los contextos sociales, económicos y culturales de cada comunidad educativa, contribuye a legitimar estructuras de exclusión. En lugar de ser herramientas para la mejora y la equidad, muchas evaluaciones terminan reforzando jerarquías preexistentes. De igual modo, Ravela (2009) propone que las prácticas evaluativas deben orientarse a procesos de retroalimentación dialógica, que permitan al profesorado y al estudiantado interpretar sus avances, tomar decisiones pedagógicas informadas y visibilizar el aprendizaje más allá de los resultados estandarizados. Esta perspectiva enfatiza la necesidad de construir una evaluación que no solo mida, sino que también enseñe, acompañe y emancipe.

Murillo y Román (2010) enfatizan que uno de los principales desafíos que enfrentan los sistemas de evaluación educativa en América Latina es dejar de concebir la evaluación como un fin en sí mismo. La consolidación de modelos centrados en la medición de resultados, el control del desempeño y la generación de rankings ha desplazado el sentido pedagógico y formativo de la evaluación. Frente a este escenario, los autores proponen una transformación sustantiva que permita convertir la evaluación en una herramienta al servicio de la mejora

auténtica, basada en principios de transparencia, inclusión y participación. Esta propuesta se alinea con una pedagogía del compromiso, entendida como una práctica crítica y situada que reconoce a docentes y estudiantes como sujetos activos en los procesos de enseñanza y evaluación, y que busca construir una educación pública más humana, justa y democrática.

Esta visión crítica de la evaluación es compartida por Zacarias (2018), quien advierte que el predominio de enfoques tecnocráticos ha limitado el potencial transformador de la evaluación al centrarla en la eficiencia y la estandarización. Desde su perspectiva, es indispensable avanzar hacia modelos evaluativos articulados con políticas educativas sustantivas, que reconozcan la diversidad, promuevan la equidad y fortalezcan la cultura profesional docente. En la misma línea, Ramírez-Quintero y Rojas-Gámez (2025) plantean la necesidad de reconceptualizar la evaluación como un proceso dialógico, contextualizado e inclusivo, que se vincule con el desarrollo integral del estudiantado y el reconocimiento de saberes diversos en el aula.

En este sentido, Ravela et al. (2016) insisten en que el gran desafío no está en seguir perfeccionando sistemas de medición, sino en restituir a la evaluación su función pedagógica, orientada al acompañamiento, la retroalimentación y la mejora de los aprendizajes. Este enfoque requiere políticas coherentes, marcos curriculares integradores y una resignificación del rol docente, entendiendo que una evaluación justa no solo mide, sino también forma, orienta y transforma.

En síntesis, la evaluación educativa en América Latina se encuentra en una tensión constante entre dos racionalidades opuestas: una lógica dominante, de corte tecnocrático, centrada en la clasificación y la vigilancia del rendimiento; y una lógica emergente, que impulsa modelos éticos, inclusivos, contextuales y centrados en el aprendizaje. Esta disputa no solo expresa diferencias técnicas o metodológicas, sino que representa visiones contrapuestas sobre el rol de la escuela, el valor del conocimiento y el sentido profundo de la educación pública en la región. El desafío para los próximos años no es menor: construir sistemas de evaluación que, en lugar de reproducir desigualdades, contribuyan activamente a democratizar el saber y a garantizar el derecho a una educación significativa y transformadora para todos y todas.

2.5.2. Evaluación educativa en Chile

La evaluación educativa en Chile ha transitado por un proceso histórico complejo, marcado por reformas estructurales, debates políticos y permanentes tensiones entre enfoques técnicos y pedagógicos. Desde fines del siglo XIX, el Estado chileno ha recurrido sistemáticamente a exámenes y pruebas como mecanismos para regular la educación, primero con fines de control centralizado y posteriormente como instrumentos de gobernanza educativa en contextos neoliberales (Falabella y Ramos, 2019). Como se aprecia en la tabla 4, en donde se incluyen en orden temporal diversas propuestas normativas que representan la evolución de la evaluación en Chile, en donde esta trayectoria alcanzó un punto de inflexión con la creación del SIMCE en 1988, consolidando un modelo evaluativo centrado en pruebas estandarizadas de alto impacto que han moldeado profundamente las dinámicas escolares y las prácticas pedagógicas.

Tabla 4. Síntesis de normativas asociadas a la evaluación educativa chilena.

Año y en-lace	Normativa / Ley / Política educativa	Enfoque principal
<u>1968</u>	Creación del SIMCE	Evaluación estandarizada nacional
<u>2009</u>	Actual Ley general de Educación	Sistema Educativo, Proyecto Educativo, Comunidad Educativa, Consejo Escolar, Registro Público de Sostenedores, Agencia de Calidad de la Educación
<u>2011</u>	Ley N.º 20.529 - Sistema Nacional de Aseguramiento de la Calidad de la Educación Escolar	Crea la Agencia de Calidad de la Educación.
<u>2016</u>	Ley N.º 20.903 - Sistema de Desarrollo Profesional Docente	Evaluación del desempeño docente, formación continua
<u>2018</u>	Decreto N.º 67 - Evaluación, calificación y promoción escolar	Decreto N.º 67 - Evaluación, calificación y promoción escolar
<u>2021</u>	Estándares Pedagógicos y Disciplinarios del CPEIP	Competencias docentes, evaluación pertinente y fundamentada
<u>2025</u>	Proceso de Actualización Curricular (consulta pública para 1º básico a IIº medio)	Reforma curricular integral, foco en evaluación formativa y desarrollo integral

Elaboración propia

Durante las últimas décadas, el país ha fortalecido su sistema de aseguramiento de la calidad mediante la promulgación de la Ley N.º 20.529 y la creación de la Agencia de Calidad

de la Educación. Este marco institucional ha promovido el uso sistemático de evaluaciones internas y externas, estableciendo indicadores para el logro de aprendizajes, el clima escolar y la gestión institucional, en concordancia con las recomendaciones internacionales (OCDE, 2017). No obstante, múltiples estudios han señalado que el sistema mantiene un sesgo tecnocrático, con fuerte énfasis en la rendición de cuentas, lo que ha generado efectos colaterales como el estrechamiento curricular, la presión sobre los establecimientos y el deterioro del sentido pedagógico de la evaluación (Castillo, 2021).

En el plano normativo, la LGE y el Decreto N.º 67 del año 2018 han introducido avances significativos en la redefinición del enfoque evaluativo. Este decreto establece que la evaluación escolar debe ser continua, integral y centrada en el aprendizaje, incorporando de forma explícita los principios de la evaluación formativa. Asimismo, promueve la retroalimentación oportuna y reconoce la atención a la diversidad como criterio fundamental en el proceso evaluativo (Ministerio de Educación, 2018). Sin embargo, la implementación de estas orientaciones enfrenta resistencias culturales y prácticas arraigadas, como la centralidad de la calificación numérica en la toma de decisiones pedagógicas.

En efecto, investigaciones recientes evidencian una tensión estructural entre el discurso formativo promovido por la normativa y la persistente cultura calificativa que domina en las escuelas. Rodríguez et al. (2025) advierten que la nota sigue funcionando como el principal instrumento de poder, jerarquización y control dentro del aula, relegando los propósitos pedagógicos de la evaluación a un segundo plano. Esta contradicción revela las dificultades para transformar prácticas evaluativas profundamente instaladas, a pesar de las intenciones declaradas por las políticas públicas.

En el sistema escolar chileno, las calificaciones se expresan en una escala numérica que va de 1,0 (nota mínima) a 7,0 (nota máxima), siendo 4,0 la nota mínima exigida para aprobar una asignatura. Esta escala, aplicada tanto en educación básica como media, se estructura a partir de ponderaciones definidas por cada establecimiento educacional, según su reglamento interno. Como establece el Decreto N.º 67, los criterios de promoción deben considerar los objetivos de aprendizaje, las características del estudiantado y, de ser necesario, la aplicación de medidas excepcionales que prioricen el bienestar del estudiantado (Ministerio de

Educación, 2018). No obstante, pese a los principios formativos promovidos por la normativa vigente, en la práctica, la calificación numérica sigue siendo el eje estructurante de las decisiones pedagógicas más relevantes, como la promoción, la repitencia y el acceso a beneficios estatales. Entre ellos destaca la gratuidad universitaria, cuya asignación considera directamente el promedio de notas de enseñanza media, reforzando el peso simbólico y funcional de la calificación como mecanismo de estratificación social y académica (Ministerio de Educación, 2023; Rodríguez et al., 2025). Esta dualidad entre el discurso de evaluación formativa y la persistencia del calificacionismo ha sido ampliamente documentada por diversos estudios críticos sobre la cultura escolar chilena (Zúñiga et al., 2023). Comprender el funcionamiento de esta estructura evaluativa es clave para interpretar las prácticas reales que predominan en la disciplina, así como las tensiones entre los propósitos declarados de la política educativa y los dispositivos concretos de regulación escolar (González Tapia, 2021; Rodríguez et al., 2025).

Paralelamente, la literatura ha reportado un fenómeno cada vez más extendido: la inflación de calificaciones. Investigaciones recientes advierten una desconexión creciente entre los resultados obtenidos en pruebas estandarizadas como el SIMCE y los promedios de notas registrados por los establecimientos, especialmente en el nivel secundario. Según Rodríguez et al. (2025), las notas escolares han aumentado sostenidamente durante la última década sin un correlato claro en los aprendizajes evaluados externamente. Esta situación se vincula, entre otros factores, con la presión que enfrentan los docentes por mostrar buenos resultados, en especial en el marco del SAE, donde el promedio de notas incide directamente en la postulación de los y las estudiantes a establecimientos de alta demanda (Castillo, 2021). A su vez, estudios como los de Avalos y Assael (2020) han advertido que esta presión puede inducir comportamientos de ajuste estratégico en la evaluación, generando distorsiones pedagógicas que afectan la credibilidad del sistema y su capacidad para retroalimentar el aprendizaje. Esta cultura de la calificación ha desplazado el sentido pedagógico de la evaluación, reduciendo su potencial formativo y profundizando su función de control, diferenciación y acreditación (Rodríguez et al., 2025; Zúñiga et al., 2023).

En síntesis, el sistema de calificaciones chileno continúa operando como un dispositivo central de estructuración escolar, que tensiona los principios formativos promovidos desde

la normativa con una cultura profundamente arraigada de cuantificación, selectividad y rendimiento. Esta contradicción plantea la necesidad urgente de repensar el lugar que ocupa la nota en el proceso educativo, incorporando enfoques más integradores, éticos y orientados al desarrollo del aprendizaje en su complejidad.

Paralelamente, el país ha comenzado a incorporar enfoques más comprensivos de la evaluación. Iniciativas como la Evaluación Progresiva y la Política para el Fortalecimiento de la Evaluación en el Aula, impulsadas por la Agencia de Calidad de la Educación y el Ministerio de Educación, proponen equilibrar los fines formativos y sumativos de la evaluación, reconociendo el potencial pedagógico del proceso evaluativo y su impacto en el aprendizaje (Ministerio de Educación, 2018). Sin embargo, su implementación ha sido desigual y altamente dependiente de las capacidades institucionales locales, lo que evidencia la necesidad de una mayor articulación entre políticas nacionales y realidades escolares.

En el plano de la evaluación docente, las críticas se han centrado en el uso del portafolio como único instrumento de valoración del desempeño profesional. Si bien se reconoce su potencial formativo, distintos estudios advierten que su aplicación presenta limitaciones en términos de validez contextual, sobrecarga administrativa y escasa conexión con las prácticas reales de aula (Cerón y Echeverría, 2020). Esta situación ha sido interpretada como expresión de una lógica de accountability gerencial, que prioriza la rendición de cuentas por sobre el desarrollo profesional situado y colaborativo (Galaz-Ruiz y Anaya-Torres, 2023).

Asimismo, se ha documentado una baja utilización de los resultados de evaluaciones a gran escala para abordar cuestiones de equidad, especialmente en relación con el género. Caniupán y Valenzuela (2025) constataron que, aunque los datos del SIMCE y el DIA se encuentran desagregados por sexo, la mayoría de las escuelas chilenas no los analizan ni los integran en sus decisiones pedagógicas. Esto refleja tanto una limitada formación docente en género como una débil articulación entre evaluación y justicia educativa.

En síntesis, la evaluación educativa en Chile se encuentra en una encrucijada entre un marco normativo que reconoce su valor formativo y una cultura escolar e institucional que aún responde, en gran medida, a lógicas de control, acreditación y rendimiento. El desafío consiste en avanzar hacia un modelo de evaluación integral, que articule políticas públicas,

prácticas pedagógicas y sentidos éticos, y que sea capaz de contribuir a una educación más justa, contextualizada y centrada en el aprendizaje.

2.6. La evaluación en Educación Física

2.6.1. La historia de la Educación Física y su relación con la evaluación

La EF ha sido históricamente configurada por corrientes médicas, militares, higienistas y pedagógicas, las cuales han delineado tanto sus propósitos como sus métodos de enseñanza y evaluación. Desde el siglo XIX, esta área del conocimiento fue incorporándose al currículo escolar con un énfasis en la disciplina corporal, la salud pública y el fortalecimiento físico de la población. En contextos europeos e iberoamericanos, como España o Chile, predominó una visión biomédica y funcionalista que promovió el ejercicio físico como herramienta para la regeneración del cuerpo social (Martínez, 2010; Ovalle y Briones, 2022).

En este periodo fundacional, la evaluación en EF no respondía a criterios pedagógicos sistemáticos, sino que se reducía a la observación del estado físico del estudiantado, sin intencionalidad formativa ni procesos de retroalimentación. A lo largo del siglo XX, la influencia de modelos gimnásticos europeos, el auge del deporte competitivo y el paradigma de rendimiento profundizaron esta perspectiva reduccionista. Como advierten Poblete y Moreno-Doña (2023), las prácticas evaluativas en EF han privilegiado históricamente la técnica, la eficacia y la comparación, en desmedro de dimensiones expresivas, afectivas y sociales del movimiento humano.

Estas prácticas tradicionales se anclaron en lo que varios autores han llamado una "racionalidad tecnocrática", donde la evaluación se emplea como mecanismo de control, estandarización y exclusión (Rannau, 2020). En el caso chileno, la persistencia de instrumentos centrados en el rendimiento físico, como los test de condición física usados para calificar, ha sido objeto de crítica por sus efectos emocionales adversos y su escasa pertinencia curricular.

Sin embargo, desde la segunda mitad del siglo XX y especialmente en las últimas décadas, diversos movimientos pedagógicos y reformas educativas han impulsado una transformación en los enfoques evaluativos. En España, la LOGSE (1990) representó un

punto de inflexión al introducir una visión curricular basada en competencias, que otorgaba protagonismo a la evaluación procesual y a la integralidad del aprendizaje corporal (Torrebadella-Flix, 2017).

Este giro ha sido acompañado por una creciente producción internacional que, desde una mirada crítica y propositiva, ha problematizado las prácticas heredadas. En el contexto anglosajón, múltiples estudios han documentado las tensiones entre discursos formativos y prácticas reproductivas. Por ejemplo, Killian y Woods (2021) evidencian que en Estados Unidos, pese a un discurso creciente en favor de la evaluación para el aprendizaje, las prácticas reales en educación escolar aún se centran en la medición del rendimiento físico y la asignación de calificaciones, con escasa integración de criterios formativos.

De manera similar, Borghouts et al. (2017) analizaron el caso de Países Bajos, mostrando que la calidad de las prácticas evaluativas en secundaria sigue siendo baja, con limitada claridad en los propósitos y escasa participación del estudiantado. En el ámbito de la formación docente, Starck et al. (2018) proponen un marco conceptual de "alfabetización evaluativa" que permita a futuros profesores comprender la evaluación no solo como técnica, sino como una práctica pedagógica crítica.

Estas propuestas dialogan con los planteamientos de Dinan-Thompson y Penney (2015), quienes sostienen que, en la educación primaria, la falta de formación específica sobre evaluación limita la implementación de enfoques inclusivos y auténticos. Por su parte, Chng y Lund (2018) abogan por un enfoque de "assessment for learning" en EF, centrado en la retroalimentación constructiva, la autoevaluación y la reflexión, claves para un aprendizaje significativo. Esta misma preocupación es retomada por Sundaresan et al. (2017), quienes denuncian la marginalidad de la evaluación de calidad en las reformas curriculares, a pesar de los discursos sobre formación integral. Finalmente, Towner et al. (2025) destacan, mediante un estudio bibliométrico, la creciente relevancia de la investigación en evaluación en EF, aunque advierten una desconexión persistente entre teoría y práctica docente.

En síntesis, la evolución histórica de la evaluación en EF da cuenta de una tensión persistente entre tradición y renovación. Mientras persisten prácticas asociadas a la medición del rendimiento y la reproducción de modelos excluyentes, emergen con fuerza enfoques que

conciben la evaluación como un acto pedagógico, ético y situado. Esta transformación no es lineal ni homogénea: implica disputas epistemológicas, resistencias institucionales y desafíos en la formación inicial y continua del profesorado. Replantear la evaluación desde esta perspectiva crítica requiere no solo revisar los instrumentos o técnicas, sino interrogar los sentidos, fines y valores que históricamente han sustentado la EF escolar.

2.6.2. Paradigmas y enfoques de evaluación en Educación Física

La evaluación en EF ha sido históricamente dominada por un enfoque técnico, vinculado a la medición objetiva del rendimiento físico, al control y a la asignación de calificaciones como fin último del proceso evaluativo (Calatayud, 2021; Mora, 2004). Desde esta lógica, evaluación y calificación han sido concebidas casi como sinónimos (Hortigüela-Alcalá et al., 2021), propiciando prácticas centradas en la administración de pruebas estandarizadas y test de condición física, que reproducen esquemas de control y jerarquización del estudiantado (Atienza et al., 2018; Sánchez-Hernández et al., 2022). Este paradigma, claramente asociado a una racionalidad instrumental y tecnocrática, ha promovido una visión de la evaluación orientada al producto y desvinculada de los procesos formativos (Oliver-Álvarez y Martos-García, 2023; Tinning, 2017).

No obstante, diversos estudios han cuestionado este modelo hegemónico por sus efectos negativos en la experiencia escolar del alumnado, especialmente de quienes presentan menor habilidad motriz o escasa afinidad con la actividad física. La aplicación de pruebas físicas estandarizadas desde un enfoque técnico no solo expone a comparaciones constantes, sino que también puede provocar ansiedad, estrés y sentimientos de exclusión entre quienes no logran los estándares exigidos (Danthony et al., 2020; Huhtiniemi et al., 2021; Monforte y Pérez-Samaniego, 2017). Estas emociones adversas se intensifican en estudiantes menos habilidosos o con cuerpos no normativos, quienes suelen experimentar frustración y rechazo hacia la asignatura, erosionando su autoestima y motivación (Beltrán-Carrillo y Devís, 2019; Oliver-Álvarez y Martos-García, 2023). Algunos estudios han descrito cómo estas prácticas generan vivencias de miedo al ridículo, temor al juicio del grupo, y hasta odio hacia la Educación Física, asociando la evaluación con situaciones emocionalmente demoledoras (Oliver et al., 2021; Wrench y Garrett, 2008). Estas consecuencias afectan también la percepción futura del alumnado sobre la actividad física, lo que puede tener implicancias a

largo plazo en su adhesión a estilos de vida activos (Camacho-Miñano y Prat, 2018). A pesar del creciente cuerpo de evidencia crítica, estas prácticas siguen profundamente arraigadas: se estima que un 90% del profesorado continúa utilizando test físicos como principal medio de evaluación en la asignatura (López-Pastor et al., 2012), perpetuando una lógica jerárquica que prioriza el rendimiento sobre el aprendizaje inclusivo y significativo (Véliz et al., 2020; Tinning, 2017).

Como alternativa al enfoque técnico y punitivo, ha emergido con fuerza un paradigma evaluativo más amplio, dialógico y situado, sustentado en marcos prácticos y críticos. Este paradigma reconoce la evaluación en EF como un proceso profundamente formativo, ético y emancipador, que busca transformar la cultura evaluativa desde una lógica de control hacia una lógica de acompañamiento y aprendizaje (Añazco-Martínez et al., 2024; Atienza et al., 2018; McArthur, 2019). En este enfoque, la evaluación deja de ser un acto finalista para convertirse en una práctica continua e integrada al proceso pedagógico, en la que la retroalimentación oportuna y la participación activa del estudiantado son ejes centrales (Herrero-González et al., 2023; Hortigüela-Alcalá et al., 2019). Se reconoce, además, que evaluar es un acto comunicativo que implica establecer acuerdos, construir criterios compartidos y generar condiciones para la autorregulación del aprendizaje.

Dentro de esta corriente crítica se han desarrollado múltiples modelos que promueven la implicación del alumnado en su proceso formativo, como la EFyC, que fomenta la coevaluación, la autoevaluación y la reflexión metacognitiva (López-Pastor, 2009; Gallardo-Fuentes et al., 2017), o la Evaluación Auténtica, que sitúa el aprendizaje en contextos significativos y reales, vinculando el saber escolar con la experiencia vivida. A estas propuestas se suma la EpP, concepto emergente que busca resignificar la evaluación en clave emancipadora, promoviendo procesos de toma de conciencia, transformación personal y responsabilidad colectiva en los espacios educativos (Sarni-Muñoz y Corbo-Bruno, 2020). Este enfoque, además de desplazar el foco desde la calificación hacia el aprendizaje, plantea una crítica profunda a los dispositivos estandarizados, proponiendo prácticas contextualizadas, justas y sensibles a la diversidad del alumnado.

Estas prácticas alternativas, si bien aún minoritarias, se han fortalecido particularmente en la formación inicial docente (Herrero-González et al., 2023; Hortigüela-Alcalá et al., 2021), donde experiencias como la coevaluación, la autoevaluación y la participación activa en la definición de criterios han mostrado un impacto positivo en la construcción de aprendizajes significativos y en la percepción del rol evaluador (Muñoz-Sepúlveda et al., 2025). Sin embargo, la transferencia de estos enfoques a los contextos escolares reales enfrenta obstáculos como la sobrecarga laboral, la presión por resultados y la resistencia a modificar prácticas arraigadas (Barrientos et al., 2023).

En síntesis, la evaluación en EF se encuentra profundamente tensionada entre dos modelos en disputa. Por un lado, el enfoque tradicional, inscrito en una racionalidad técnica, continúa anclado en la medición del rendimiento físico, la comparación interindividual y la lógica meritocrática que ha dominado históricamente la asignatura. Por otro, emergen propuestas alternativas que reconfiguran la evaluación como una práctica pedagógica situada, orientada a la participación activa del estudiantado, la mejora continua y la construcción de experiencias educativas más humanas y equitativas (Camacho-Miñano y Prat, 2018; Calatayud, 2021).

Esta disputa trasciende lo metodológico y evidencia una pugna de sentidos sobre el papel de la EF en la escuela: si ha de limitarse a seleccionar cuerpos aptos según criterios normativos, o si, por el contrario, debe constituirse como un espacio para el desarrollo integral, la inclusión y la justicia social (Gallardo-Fuentes et al., 2017; McArthur, 2019). Lo que está en juego no es solo cómo se evalúa, sino también qué se valora, para qué se enseña y qué tipo de ciudadanía se quiere formar. Por ello, resignificar la evaluación en EF supone un imperativo ético y político que interpela tanto a las prácticas docentes como a las políticas educativas.

2.6.3. La Educación Física en Chile y su relación con la evaluación

La historia de la EF en Chile ha estado profundamente marcada por influencias médicas, pedagógicas, políticas y culturales, reflejo de un proceso de institucionalización que ha oscilado entre el control corporal, la promoción de la salud y la formación integral. Durante las primeras décadas del siglo XX, la EF se vinculó estrechamente a discursos higienistas y

a la eugenesia, promoviendo prácticas centradas en la disciplina corporal y la medición estandarizada del cuerpo, especialmente a través de pruebas de fuerza, flexibilidad y resistencia (Pérez Ramírez, 1993; Ovalle Letelier y Briones Molina, 2022). Este enfoque reduccionista consolidó una visión instrumental del cuerpo, priorizando el rendimiento físico por sobre los procesos de aprendizaje o las dimensiones afectivas y sociales del movimiento (Poblete Gálvez et al., 2014).

La Revista Chilena de EF de los años treinta ejemplifica esta orientación, al difundir una formación física centrada en la mejora de la condición corporal como base para el bienestar nacional, excluyendo visiones pedagógicas o críticas sobre el valor educativo del movimiento (Ovalle Letelier y Briones Molina, 2022). No obstante, hacia las últimas décadas del siglo XX, comenzó a gestarse una transformación curricular progresiva que culminó en la década del 2000 con un cambio significativo: la asignatura pasó de denominarse "Educación Física" a "Educación Física y Salud", primero con el Marco Curricular de 2009 para 1.º a 6.º básico y luego en 2012 para 7.º básico a 2.º medio. Este giro nominal fue también conceptual, integrando en su propósito la formación valórica, el bienestar emocional y la promoción de estilos de vida saludables (Ministerio de Educación, 2012).

En paralelo a esta ampliación curricular, la evaluación en EF fue adquiriendo nuevos sentidos. Tradicionalmente dominada por la lógica tecnocrática de la medición y la clasificación, ha enfrentado tensiones y cuestionamientos, especialmente con la promulgación del Decreto N.º 67 en 2018, que estableció orientaciones para una evaluación formativa, integral y continua. Este decreto enfatizó la importancia de la retroalimentación, el seguimiento de procesos y la adecuación de los instrumentos evaluativos al contexto y a la diversidad del estudiantado (Ministerio de Educación, 2018).

Sin embargo, este avance normativo ha coexistido con políticas contradictorias, como la implementación del SIMCE de EF en 2010, el cual se centró en variables biométricas como el índice de masa corporal o la resistencia aeróbica, evidenciando una tensión entre el discurso pedagógico inclusivo y las prácticas evaluativas estandarizadas. Investigaciones críticas han señalado que esta versión del SIMCE reforzó una visión biologicista del cuerpo y de la evaluación, deslegitimando el rol del profesorado y debilitando el potencial formativo

de la asignatura (Luarte-Rocha et al., 2016; Moreno et al., 2014). Más recientemente, el anuncio de una nueva versión piloto del SIMCE de EF para el año 2025, con un conjunto de test físicos actualizados, reactiva este debate y evidencia la necesidad de reflexionar sobre el tipo de evaluación que se desea promover desde la política pública.

Por otro lado, diversos estudios han evidenciado que, pese a las orientaciones normativas actuales, como las Bases Curriculares de 2016 que prohíben calificar mediante test de condición física, estas pruebas continúan utilizándose en la práctica, reforzando lógicas meritocráticas y excluyentes (Castillo-Retamal et al., 2024; Portes-Junior et al., 2022). Esta desconexión entre las orientaciones oficiales y la práctica docente constituye una de las principales tensiones actuales de la evaluación en EF chilena.

Afortunadamente, también emergen iniciativas transformadoras. Experiencias desarrolladas en la formación docente, como las impulsadas por Gallardo-Fuentes et al. (2017, 2019, 2022) y Herrero-González et al. (2023), han demostrado el potencial de la EFyC para resignificar la práctica evaluativa. Estas propuestas promueven la participación activa del estudiantado, la retroalimentación significativa y el uso de rúbricas contextualizadas, favoreciendo una cultura pedagógica más ética, reflexiva y centrada en el aprendizaje. Asimismo, investigaciones aplicadas a contextos como la gimnasia artística, las danzas tradicionales o los deportes colectivos han mostrado que la incorporación de criterios diversificados y el uso de instrumentos adaptativos pueden mejorar no solo el aprendizaje, sino también la motivación y la percepción de justicia educativa entre el estudiantado (Añazco-Martínez et al., 2024; Chihuailaf-Vera et al., 2024; Muñoz-Sepúlveda et al., 2025b).

Finalmente, el artículo de Mujica Johnson et al. (2023) aporta una mirada crítica al señalar las tensiones epistemológicas presentes en los estándares disciplinares para la formación docente en EF. La exclusión de conceptos claves como "actividad física", que sí están presentes en el currículo escolar, revela un desfase entre las orientaciones formativas y las demandas de la educación básica y media. Estas omisiones, según los autores, no solo empobrecen el debate académico, sino que también debilitan la coherencia del sistema educativo en su conjunto.

En suma, la EF en Chile y su relación con la evaluación se encuentra en una fase de transición. Aunque persisten prácticas tradicionales marcadas por la lógica del rendimiento y la medición, emergen también propuestas pedagógicas críticas que abogan por una evaluación situada, inclusiva y coherente con los principios de formación integral. El proceso de actualización curricular en curso, junto con la implementación del nuevo SIMCE, representa una oportunidad crucial para reorientar estas tensiones y construir una evaluación verdaderamente al servicio del aprendizaje y del bienestar estudiantil.

2.6.4. La formación del profesorado en Educación Física

En el contexto chileno, la formación inicial del profesorado de EF se encuentra regulada por diversas políticas públicas que buscan asegurar la calidad, pertinencia y equidad de los programas formativos. Entre estas normativas destacan la Ley N.º 20.903, que crea el Sistema de Desarrollo Profesional Docente (Ministerio de Educación, 2016), el MBE y los Estándares Pedagógicos y Disciplinarios definidos por el CPEIP.

El MBE establece cuatro dominios esenciales para la práctica docente: preparación de la enseñanza, creación de ambientes propicios para el aprendizaje, enseñanza para el aprendizaje de todos los y las estudiantes y compromiso profesional. En todos ellos, la evaluación se concibe como una herramienta fundamental para la mejora continua de los aprendizajes (Ministerio de Educación, 2021). A esto se suman los Estándares Disciplinarios para la Pedagogía en EF, los cuales explicitan que los futuros docentes deben ser capaces de diseñar instrumentos evaluativos contextualizados, implementar procedimientos variados y utilizar los resultados para retroalimentar el proceso educativo (CPEIP, 2021).

Sin embargo, a pesar de este marco normativo, diversas investigaciones han evidenciado una brecha entre las orientaciones institucionales y las prácticas reales de formación y enseñanza. Estudios como los de Merellano-Navarro et al. (2019) y Muñoz-Sepúlveda et al. (2024) indican que la evaluación en la formación inicial sigue centrada en lógicas tradicionales, priorizando el rendimiento físico sobre aspectos formativos o socioemocionales. Asimismo, se ha documentado una escasa articulación entre los contenidos teóricos y las prácticas pedagógicas desarrolladas en los centros escolares.

A esto se suma la preocupación por la baja presencia de contenidos asociados a la inclusión educativa y la equidad de género en los programas de pedagogía en EF. Un análisis de mallas curriculares reveló que menos del 6% de las horas de formación estaban destinadas a temáticas de diversidad, lo que limita las capacidades del profesorado para generar ambientes de aprendizaje democráticos y respetuosos de la diferencia (Chihuailaf-Vera et al., 2024; Castillo-Retamal et al., 2021).

No obstante, algunas experiencias innovadoras comienzan a marcar una ruta de transformación. Investigaciones desarrolladas en el contexto universitario chileno han documentado los beneficios de incorporar la EFyC, las rúbricas participativas y la coevaluación como estrategias para potenciar la comprensión pedagógica del futuro profesorado y mejorar sus competencias evaluativas (Gallardo-Fuentes et al., 2017; Muñoz-Sepúlveda y Valencia-peris, 2022).

Asimismo, la formación continua ha sido reforzada como eje del desarrollo profesional, especialmente a partir de la Ley N.º 20.903, que reconoce formalmente la carrera docente y promueve oportunidades de perfeccionamiento. Programas impulsados por el CPEIP y universidades acreditadas permiten a docentes en ejercicio actualizar sus prácticas, incorporar nuevas metodologías y responder a los desafíos de una educación inclusiva y de calidad. Los Estándares Pedagógicos y Disciplinarios han ido incorporando paulatinamente una visión más compleja de la evaluación. En la versión de 2014, el Estándar 6 se centraba en aspectos operativos como la planificación, aplicación y análisis de instrumentos evaluativos. En cambio, el Estándar C de 2021 enfatiza el carácter ético, situado y formativo de la evaluación, promoviendo el uso de instrumentos pertinentes, coherentes con los propósitos formativos y adaptados a la diversidad del estudiantado (CPEIP, 2014, 2021). Este tránsito revela una evolución conceptual importante, aunque aún convive con prácticas heredadas del paradigma técnico, muchas veces impulsadas por presiones institucionales o por la falta de sistematización en la formación práctica (Barrientos et al., 2023; Mujica Johnson et al., 2023).

En definitiva, la formación inicial del profesorado de EF en Chile se encuentra en un proceso de transición. Aunque los marcos normativos ofrecen una base sólida para el

desarrollo de competencias evaluativas pertinentes, inclusivas y pedagógicamente fundamentadas, su implementación efectiva exige un compromiso institucional sostenido, la mejora continua de los programas formativos y una mayor articulación entre la teoría, la práctica y las experiencias previas del estudiantado que se forma como futuro profesorado.

2.7. Estado de la cuestión

La evaluación en la asignatura de EF ha cobrado creciente protagonismo en la literatura científica internacional, al constituirse como un componente clave en la experiencia escolar del estudiantado y en la construcción de significados en torno al cuerpo, la motricidad y el aprendizaje. Esta atención se explica, en gran parte, por las tensiones persistentes entre enfoques evaluativos tradicionales, de corte técnico y reproductivo, y propuestas emergentes que abogan por una evaluación más formativa, reflexiva e inclusiva (Herrero-González et al., 2023; López-Pastor et al., 2020). En el ámbito latinoamericano, esta disputa ha adoptado formas concretas que reflejan tanto las herencias de modelos normalizadores como los esfuerzos por resignificar la evaluación desde marcos pedagógicos más democráticos. Secchi et al. (2016) y Stieg y dos Santos (2021) identifican con claridad dos corrientes predominantes: una, centrada en indicadores cuantificables como el rendimiento físico, la condición cardiorrespiratoria y la salud, vinculada a una racionalidad tecnocrática; y otra, de carácter cualitativo, que se articula en torno a procesos formativos, la reflexión docente, la participación activa del alumnado y la atención a la diversidad.

En este escenario, diversos estudios han subrayado que el enfoque cuantitativo dominante, legitimado por sistemas estandarizados de evaluación como las pruebas físicas, tiende a reproducir dinámicas excluyentes, jerárquicas y descontextualizadas (Moreno-Doña et al., 2014; Oliver-Álvarez y Martos-García, 2023), mientras que los enfoques alternativos han sido asociados con un mayor potencial para fomentar el desarrollo integral, la equidad y la motivación (Cañadas et al., 2021; Barba-Martín et al., 2020). Estas tensiones no son meramente metodológicas, sino que reflejan visiones contrapuestas sobre los fines educativos de la EF y el tipo de ciudadanía corporal que se desea promover.

En el caso chileno, el enfoque técnico ha predominado históricamente en la evaluación de la asignatura de EFyS, consolidándose institucionalmente a través de políticas públicas

como el SIMCE. Desde su implementación, el SIMCE para EF ha priorizado la evaluación de variables objetivas como el rendimiento muscular, la resistencia cardiorrespiratoria, la flexibilidad y los indicadores antropométricos (Ministerio de Educación, 2010), con una lógica centrada en la estandarización, la comparación interindividual y la clasificación. Esta orientación ha sido fuertemente cuestionada por su reduccionismo pedagógico, su escaso aporte al aprendizaje significativo y su incompatibilidad con enfoques inclusivos y formativos (Moreno-Doña et al., 2014; Castillo-Retamal et al., 2024). Es relevante comentar que para el año 2025 se pretende volver a retomar las evaluaciones del SIMCE en la asignatura de EFyS, con baterías de pruebas físicas.

Las investigaciones recientes coinciden en advertir los efectos adversos de las prácticas evaluativas centradas en pruebas físicas estandarizadas, particularmente en estudiantes que no logran cumplir con los estándares exigidos. Estas prácticas han sido asociadas a experiencias de ansiedad, vergüenza, exclusión y rechazo hacia la asignatura, afectando negativamente la motivación y el vínculo con la actividad física a largo plazo (Muñoz-Sepúlveda et al., 2025; Oliver-Álvarez y Martos-García, 2023; Corbin, 2010; Camacho-Miñano y Girela-Rejón, 2017; Danthony et al., 2020). Aunque las Bases Curriculares de EFyS desaconsejan utilizar los test de condición física con fines calificativos (Ministerio de Educación, 2016), su persistente uso en las escuelas revela una desconexión entre las orientaciones normativas y la práctica docente.

Varios estudios han evidenciado cómo estas formas tradicionales de evaluación refuerzan jerarquías corporales, prácticas discriminatorias por género y patrones de gordofobia, lo que limita la participación de estudiantes con menor habilidad motriz o características físicas alejadas del ideal normativo (Alvariñas-Villaverde y Pazos-González, 2018; Chihuailaf-Vera et al., 2024; Devís-Coret y Martos-García, 2023). Frente a este panorama, emergen con fuerza enfoques evaluativos alternativos, como la EFyC, la evaluación auténtica y la evaluación por competencias que promueven la retroalimentación continua, la participación activa del estudiantado y la contextualización pedagógica del proceso evaluativo (Bores-García et al., 2021; Cañadas et al., 2021; Herrero-González et al., 2023; Hortigüela-Alcalá et al., 2015; Sarni-Muñiz y Corbo-Bruno, 2020).

Estos enfoques han demostrado contribuir al fortalecimiento de la motivación intrínseca, la autoeficacia y el compromiso del estudiantado con la asignatura (González-Cutre y Sicilia, 2019; Leenknecht et al., 2020), así como a la transformación de las concepciones evaluativas del profesorado en formación (Gallardo-Fuentes et al., 2017; Muñoz-Sepúlveda et al., 2024). No obstante, su implementación enfrenta múltiples obstáculos, entre ellos, la débil articulación entre los planes formativos universitarios y las prácticas escolares reales (Castillo-Retamal et al., 2022; Merellano-Navarro et al., 2019), así como la carga laboral, la burocracia y la escasez de espacios para la reflexión docente.

En este contexto, si bien el Decreto N.º 67 (Ministerio de Educación, 2018) promueve una evaluación integral y orientada al aprendizaje, la realidad muestra una hegemonía de lógicas sumativas, centradas en la calificación numérica como principal mecanismo de validación. Esta brecha entre el discurso normativo y la práctica concreta evidencia la urgencia de revisar las condiciones estructurales, institucionales y formativas que limitan el despliegue de una evaluación verdaderamente transformadora (López-Pastor, 2005; Santos Guerra, 1995). A pesar de los avances conceptuales en torno a enfoques evaluativos más formativos e inclusivos, se constata una notable ausencia de investigaciones empíricas que aborden específicamente las calificaciones en la asignatura de Educación Física en el contexto chileno. Esta laguna resulta especialmente significativa considerando que el Decreto N.º 67 del Ministerio de Educación (2018) faculta a cada establecimiento educacional a elaborar su propio reglamento de evaluación, en función de las orientaciones generales establecidas por la normativa nacional. Dada la heterogeneidad del sistema escolar chileno caracterizado por su geografía extensa y diversidad sociocultural, surge una pregunta crucial: ¿existen tendencias homogéneas o divergentes en los sistemas de calificación de EF a lo largo del país desde la implementación de este decreto? Esta interrogante no ha sido suficientemente explorada en la literatura, lo que refuerza la necesidad de estudios que analicen no solo los discursos y percepciones sobre la evaluación, sino también las formas concretas en que esta se operacionaliza en términos de calificación, sus criterios, instrumentos y consecuencias pedagógicas. La presente investigación se sitúa precisamente en ese vacío, con el propósito de aportar evidencia situada y crítica sobre las prácticas evaluativas y calificadoras en EFyS en el contexto chileno contemporáneo.

En síntesis, el panorama actual de la evaluación en EF se encuentra marcado por una tensión estructural entre la persistencia de modelos técnicos y reproductivos centrados en el rendimiento físico, y el avance, aún incipiente de enfoques formativos, dialógicos y emancipadores. A pesar de los desarrollos teóricos, normativos y metodológicos que promueven una evaluación más justa, inclusiva y significativa, su implementación en las aulas continúa siendo fragmentada y condicionada por factores estructurales, culturales y formativos. Esta brecha entre el discurso y la práctica evidencia la necesidad de revisar críticamente la formación inicial docente, fortalecer la articulación entre política educativa y realidad escolar, y promover una cultura evaluativa que supere la lógica meramente calificadora para convertirse en una herramienta de aprendizaje y justicia pedagógica. La presente investigación se inscribe en este escenario, con el objetivo de aportar evidencia empírica que contribuya a comprender y transformar las prácticas evaluativas en Educación Física desde una mirada situada, crítica y transformadora.

2.7.1. Investigaciones en evaluación desde la perspectiva del estudiantado

La evaluación en EF ha sido objeto de creciente atención desde la perspectiva del estudiantado, especialmente por su potencial influencia en variables motivacionales, afectivas y formativas (Pérez-Samaniego et al., 2011). Comprender cómo el alumnado vive y valora estas prácticas resulta clave para analizar su impacto en la motivación, la adherencia a la actividad física y la construcción de actitudes hacia la asignatura. Lejos de ser una cuestión accesorio, estas percepciones ofrecen pistas fundamentales sobre la calidad pedagógica del proceso evaluativo y sus efectos en la trayectoria educativa y corporal del estudiantado (Atienza et al., 2018; Fulano-Vargas y Meneses-Runza, 2023).

La evidencia acumulada, tanto en esta investigación como en estudios previos, ha mostrado que los procedimientos evaluativos tradicionales, centrados en el rendimiento físico, las pruebas estandarizadas y la cuantificación de resultados, tienden a generar experiencias negativas, especialmente en estudiantes con menor competencia motriz o alejados del ideal corporal dominante (Beltrán-Carrillo y Devís, 2019; Danthony et al., 2020; Huhtiniemi et al., 2021). En Chile, por ejemplo, estas prácticas han sido asociadas a sentimientos de ansiedad, frustración, inseguridad y rechazo hacia la asignatura, particularmente cuando evaluaciones como el test de Course Navette se emplean como único

criterio de calificación (Castillo-Retamal et al., 2024; Corbin, 2010; López-Gil et al., 2020; Oliver et al., 2021).

Los relatos estudiantiles recogidos en esta tesis refuerzan estos hallazgos. Se reportaron recuerdos de malestar físico y emocional, escasa claridad pedagógica y sensación de exposición pública al ser evaluados bajo parámetros rígidos o competitivos. En este contexto, la evaluación es percibida como una instancia de presión más que como una oportunidad de aprendizaje. También se identificaron prácticas punitivas, como el uso de la nota para premiar o castigar, las cuales fomentan una cultura jerárquica de la evaluación y refuerzan desigualdades estructurales (Curilem et al., 2012; Devís-Coret y Martos-García, 2023; Chihuailaf-Vera et al., 2024).

Otro elemento relevante es el impacto de la evaluación en la reproducción de estereotipos de género, gordofobia y exclusiones corporales. Diversos estudios han advertido que los estándares normativos de rendimiento o estética afectan diferencialmente a mujeres y estudiantes con sobrepeso u obesidad, generando vivencias de vergüenza y autoexclusión (Alvariñas-Villaverde y Pazos-González, 2018; Pastor-Vicedo et al., 2019; Scharagrodsky, 2004). Esta investigación confirmó esas discriminaciones, sobre todo en contextos donde se mide el peso corporal frente al grupo o se enfatizan tareas que privilegian cierto tipo de cuerpos.

En contraposición, cuando la evaluación se realiza desde enfoques formativos y participativos, las percepciones estudiantiles cambian radicalmente. La inclusión de criterios claros, la posibilidad de autoevaluación y coevaluación, y la entrega de retroalimentación oportuna son altamente valoradas por el estudiantado, ya que promueven la comprensión, la implicación y la equidad (Cañadas et al., 2021; Galán y López, 2012; Hortigüela-Alcalá et al., 2015). Estas prácticas potencian la motivación intrínseca, el disfrute por el aprendizaje y la construcción de una autoimagen positiva en torno a la actividad física (González-Cutre y Sicilia, 2019; Leenknecht et al., 2021; Mackley et al., 2022; Sevil-Serrano et al., 2020).

Además, estas percepciones están mediadas por la formación inicial docente. Investigaciones desarrolladas en universidades chilenas muestran que, cuando el profesorado en formación es expuesto a modelos alternativos, como la evaluación formativa, la

coevaluación o el uso reflexivo de rúbricas, tienden a valorar más estas estrategias y a proyectarlas en sus futuras prácticas profesionales (Gallardo-Fuentes et al., 2017; Gallardo et al., 2022). Incluso, en asignaturas específicas como gimnasia artística, se ha demostrado que adaptar criterios e instrumentos facilita tanto el aprendizaje como la percepción de justicia evaluativa (Muñoz-Sepúlveda et al., 2025).

En síntesis, las percepciones del estudiantado sobre la evaluación en EF oscilan entre el rechazo y la valoración, en función del enfoque adoptado. Estas vivencias no solo permiten comprender los efectos inmediatos de las prácticas evaluativas, sino que también ofrecen orientaciones valiosas para transformar la cultura evaluativa de la asignatura hacia modelos más éticos, inclusivos y pedagógicamente significativos.

2.8.2. Investigaciones sobre las concepciones y prácticas del profesorado

La perspectiva del profesorado, tanto en formación como en ejercicio, ha sido objeto de múltiples estudios recientes que buscan comprender las creencias, prácticas y tensiones que configuran su forma de evaluar en la asignatura de EF. En general, la literatura evidencia una alta prevalencia de prácticas evaluativas tradicionales, centradas en la medición del rendimiento físico, el uso de test estandarizados y la calificación como fin último del proceso (Hurtado-Guerrero et al., 2024; López-Pastor et al., 2012). Esta hegemonía del paradigma técnico responde, en parte, a la persistencia de una racionalidad instrumental que prioriza la cuantificación y la homogeneidad, incluso cuando existen marcos normativos que promueven modelos más formativos y dialógicos.

En el contexto chileno, estudios como el de Portes-Junior et al. (2022) han evidenciado una fuerte presencia de pruebas de condición física como herramienta principal de evaluación en EF en enseñanza media. Estas prácticas, aunque extendidas, generan preocupación debido a sus efectos negativos sobre el bienestar del estudiantado, y reflejan una desconexión entre los marcos curriculares y la práctica real en las aulas.

Complementariamente, investigaciones como la de Castillo-Retamal et al. (2024) profundizan en las percepciones del profesorado en ejercicio, revelando una tensión entre la intención pedagógica y las condiciones estructurales del sistema escolar chileno. Entre las principales dificultades se encuentran la falta de tiempo, la escasa formación específica en

evaluación, y una cultura institucional que refuerza la lógica sumativa por sobre enfoques centrados en el aprendizaje.

No obstante, también se reportan experiencias exitosas y transformadoras. En universidades chilenas, diversos estudios han mostrado cómo la implementación de modelos como la EFyC puede modificar las concepciones evaluativas del profesorado en formación, promoviendo prácticas más inclusivas, participativas y coherentes con una visión crítica de la enseñanza (Gallardo-Fuentes et al., 2017; Gallardo et al., 2022). Estas propuestas han demostrado fortalecer competencias docentes como la capacidad de diseñar instrumentos pertinentes, proporcionar retroalimentación significativa y adaptar los criterios de evaluación a la diversidad del estudiantado.

En suma, el cuerpo de evidencia científica señala que, aunque el profesorado de EF en Chile y en otros contextos continúa anclado mayoritariamente a prácticas tradicionales, existen aperturas importantes hacia la innovación pedagógica, especialmente cuando la formación inicial y continua se articula con modelos evaluativos formativos, reflexivos y situados en la realidad del aula.

2.8.3. Desafíos actuales en la investigación sobre evaluación en Educación Física

En la última década, la investigación en evaluación de la EF ha experimentado un crecimiento sostenido, aunque no exento de desafíos metodológicos y epistemológicos. Una de las principales dificultades radica en la persistencia de diseños descriptivos y enfoques cuantitativos que, si bien aportan datos relevantes sobre tendencias generales, limitan la comprensión profunda de las experiencias subjetivas del profesorado y del estudiantado en torno a la evaluación (Danthony et al., 2020; López-Pastor et al., 2020). Diversos estudios han señalado que la hegemonía de instrumentos como los test físicos continúa acaparando el interés investigativo, desplazando líneas más cualitativas o participativas que permitan comprender los efectos emocionales, motivacionales y pedagógicos de las prácticas evaluativas (Beltrán-Carrillo y Devís, 2019; Muñoz-Sepúlveda et al., 2025b). A ello se suma la escasa problematización sobre el rol de las calificaciones, tema prácticamente ausente en la literatura especializada, pese a su centralidad en los sistemas escolares.

Otro de los desafíos destacados es la escasa conexión entre investigación y práctica. Numerosos trabajos evidencian una distancia significativa entre los marcos teóricos que promueven modelos inclusivos, formativos y emancipadores, y las realidades escolares, marcadas aún por racionalidades técnicas y estructuras normativas rígidas (Castillo-Retamal et al., 2022; Portes-Junior et al., 2022). En ese marco, se ha identificado como desafío recurrente la necesidad de investigaciones contextualizadas, que consideren no solo la diversidad del estudiantado, sino también las condiciones institucionales, la formación docente y las culturas evaluativas prevalentes en cada territorio.

Finalmente, la literatura advierte sobre la falta de estudios longitudinales que permitan analizar transformaciones sostenidas en las prácticas evaluativas, así como la necesidad de promover metodologías colaborativas (investigación-acción, estudios narrativos, análisis de caso) que integren activamente a los sujetos evaluados y evaluadores en el diseño de la propia investigación (Gallardo-Fuentes et al., 2022; Herrero-González et al., 2023). Superar estos vacíos permitirá no solo enriquecer la producción científica en el área, sino también contribuir a la mejora de la evaluación en Educación Física desde una perspectiva crítica y transformadora.

Otro conjunto de desafíos emergentes en la investigación reciente sobre evaluación en EF dice relación con su impacto desbalanceado entre sus dimensiones actitudinales, conceptuales y procedimentales. En primer lugar, múltiples estudios han comenzado a indagar cómo las prácticas evaluativas, especialmente aquellas centradas en el rendimiento físico, inciden directamente en la motivación del estudiantado, su autoimagen corporal y su actitud hacia la asignatura (Bores-García et al., 2021; Cañadas et al., 2021; Hortigüela-Alcalá et al., 2015). Sin embargo, estas investigaciones suelen tener un carácter exploratorio y escasa continuidad longitudinal, lo que limita una comprensión profunda y sostenida de sus implicancias en la adherencia a la práctica físico-deportiva.

En segundo lugar, existe una notoria falta de estudios que aborden el vínculo entre la evaluación y el currículum oculto en EF. La reproducción de estereotipos de género, la discriminación por condición corporal o la naturalización de jerarquías motrices a través de la evaluación constituyen formas solapadas de exclusión que aún no han sido suficientemente

problematizadas en la literatura (Alvariñas-Villaverde y Pazos-González, 2018; Chihuailaf-Vera et al., 2024; Devís-Coret y Martos-García, 2023). Este vacío teórico limita el desarrollo de propuestas evaluativas que promuevan efectivamente la justicia social en el aula de EF.

Finalmente, aunque algunos marcos normativos y estudios recientes abogan por modelos de evaluación con enfoque inclusivo, ético y situado, la producción científica sigue estando concentrada en países específicos, con escasa presencia de estudios desde contextos del Sur Global, como América Latina o África. Esta asimetría limita la posibilidad de construir un panorama verdaderamente representativo de las problemáticas evaluativas globales y refuerza una visión eurocéntrica de la evaluación (Añazco-Martínez et al., 2024; Gallardo-Fuentes et al., 2022). Se requiere, por tanto, un esfuerzo deliberado por promover investigaciones críticas, interseccionales y contextualizadas que aborden los vínculos entre evaluación, equidad y transformación educativa en la EF.



VNIVERSITAT E VALÈNCIA

CAPÍTULO 3

MARCO METODOLÓGICO

Para dar cumplimiento a los objetivos específicos de esta tesis doctoral, la investigación fue organizada en dos etapas complementarias que corresponden a distintas fuentes de información: por un lado, las experiencias del estudiantado respecto a la evaluación en EF; y, por otro, las percepciones y prácticas del profesorado en ejercicio. En ambas etapas se emplearon enfoques coherentes con el paradigma cualitativo, utilizando técnicas e instrumentos diferenciados según el tipo de participante y el propósito de cada fase.

Es fundamental que, antes de presentar el detalle de las dos etapas con sus respectivas características, se aborde el apartado del paradigma de investigación. Por ello, previo a la presentación general en la tabla 4, se desarrolla el siguiente apartado.

3.1. Paradigma de investigación

La presente investigación se enmarca predominantemente en el paradigma interpretativo, el cual se caracteriza por una orientación descriptiva y comprensiva de la realidad social, usando para ello metodología cualitativa. Este paradigma asume que los fenómenos educativos deben ser comprendidos desde las experiencias, significados y contextos de quienes los viven (Pérez-Samaniego, 2021). En este enfoque, la investigación se desarrolla en escenarios naturales y busca acceder a las voces y significados construidos por los y las actores involucrados, priorizando la subjetividad y el sentido de las vivencias educativas (Denzin y Lincoln, 2011).

El propósito principal de esta tesis es profundizar en el proceso de evaluación en Educación Física desde la experiencia del estudiantado y del profesorado en ejercicio, por lo que se demanda una aproximación interpretativa, atenta a los relatos y a los contextos escolares y formativos que configuran dichas prácticas evaluativas. Tanto en la Etapa 1, centrada en las experiencias escolares del estudiantado universitario, como en la Etapa 2, enfocada en las percepciones y prácticas del profesorado de EF, se privilegió la comprensión situada de los discursos bajo una lógica cualitativa que reconoce la construcción social del conocimiento (Goodson, 2013; Monforte y Pérez Samaniego, 2021).

No obstante, desde el inicio de la segunda etapa de la investigación, centrada en el profesorado en ejercicio, emergió de forma inductiva la necesidad de complementar los hallazgos cualitativos con un análisis de datos cuantitativos secundarios. En particular,

algunas entrevistas del Estudio 3 aludieron al uso habitual de calificaciones elevadas en la asignatura de EF, lo que motivó una nueva línea de indagación orientada a examinar el comportamiento real de dichas calificaciones a nivel nacional. Esta necesidad dio origen al Estudio 4, donde se incorporó el análisis de las bases públicas del MINEDUC entre los años 2018 y 2022. Tal decisión metodológica condujo a la adopción de un diseño secuencial exploratorio, característico de los enfoques mixtos, en el cual los hallazgos cualitativos orientan la inclusión de datos cuantitativos con el propósito de enriquecer la comprensión del fenómeno evaluativo (Creswell y Plano Clark, 2018; Greene, 2007).

Este componente cuantitativo no altera el paradigma central de la investigación, sino que amplía la comprensión del fenómeno evaluativo, permitiendo contrastar las percepciones docentes con tendencias empíricas a nivel nacional. Así, se articula una propuesta metodológica flexible y coherente con el principio de complementariedad, manteniendo la primacía de la mirada cualitativa e interpretativa, sin perder rigor ni profundidad analítica (Johnson et al., 2007).

En suma, la inclusión de un componente cuantitativo en la etapa final responde a la necesidad de triangular perspectivas, enriquecer el análisis y fortalecer la validez interpretativa de los hallazgos. Esta decisión se alinea con una visión crítica y reflexiva de la evaluación educativa, que exige enfoques metodológicos abiertos y sensibles a la complejidad del objeto de estudio.

La Tabla 5 sintetiza esta organización metodológica, permitiendo visualizar la relación entre los estudios desarrollados, los paradigmas y enfoques asumidos, así como los instrumentos empleados en cada caso.

Tabla 5. Organización metodológica de la tesis según los objetivos y etapas.

Objetivo de la investigación	Etapas	Estudio	Paradigma	Enfoque	Técnica de recolección de datos	Instrumento de recolección de datos	Fecha de recolección de datos
1. Caracterizar las experiencias vividas sobre evaluación en las clases de Educación Física escolar en estudiantes de una Facultad de Educación.	Etapas 1: Desde las experiencias y perspectivas del estudiantado.	1. Dime cuánto corres y te diré qué nota mereces. Evaluación en la Educación Física chilena.	Cualitativo	Narrativo	Relatos narrativos	Relatos escritos redactados por el estudiantado	Junio/2022
2. Indagar sobre la influencia del proceso evaluativo de la Educación Física escolar en las actitudes del alumnado hacia la asignatura.		2. Evaluación en Educación Física escolar y su impacto en la percepción de la asignatura: vivencias narrativas de estudiantes de pedagogía		Narrativo-interpretativo	Grupo focal	Guion de grupo focal	Diciembre 2022
3. Profundizar en la conceptualización y perspectivas evaluativas del profesorado de Educación Física y Salud en ejercicio.	Etapas 2. Desde las percepciones y prácticas evaluativas del profesorado de EF.	3. Del rendimiento al esfuerzo y actitud: Enfoques evaluativos en Educación Física desde la experiencia docente		Interpretativo	Entrevista y análisis documental	Guion de entrevista semiestructurada / Análisis de instrumentos de evaluación	Septiembre 2023 a enero 2024
4. Analizar las prácticas evaluativas del profesorado en ejercicio, desde sus experiencias escolares y universitarias en evaluación.		4. La nota máxima está de moda. Análisis de las prácticas evaluativas en la asignatura de Educación Física y Salud chilena.	Mixto	Secuencial exploratorio	Entrevistas semiestructuradas y análisis de bases de datos de calificaciones .	Guion de entrevista semiestructurada / Análisis de bases de datos públicas del Mineduc (2018–2022)	Septiembre 2023 a julio 2024

Elaboración propia.

3.2. Metodología etapa 1

3.2.1. Enfoque de investigación

La primera etapa de esta tesis doctoral se enfoca en los objetivos específicos 1 y 2, orientados a comprender las experiencias de evaluación en la asignatura de EF desde la perspectiva del estudiantado (ver Tabla 5). Para ello, se adopta un enfoque cualitativo bajo la tradición narrativa, el cual permite explorar en profundidad los significados atribuidos por los sujetos a sus vivencias escolares. Esta metodología se centra en el análisis de relatos personales narrativos, mediante los cuales el estudiantado expresa libremente sus percepciones, emociones y valoraciones sobre el proceso evaluativo (Bolívar, 2002; León, 2015).

La investigación narrativa se inscribe en un marco hermenéutico-interpretativo que reconoce a las personas como sujetos históricos y sociales, portadores de experiencias subjetivas que deben ser comprendidas en su contexto (Silva, 2017). Desde esta perspectiva, narrar no implica simplemente relatar hechos, sino interpretar y dar sentido a vivencias que configuran la identidad del sujeto, en este caso, sobre las vivencias y recuerdos en evaluación de la asignatura de EF en edad escolar. Así, los relatos se constituyen en vehículos de comprensión profunda de los procesos educativos, revelando no solo las percepciones inmediatas del estudiantado, sino también sus tramas personales, memorias escolares y relaciones con otros significativos, como docentes o compañeros (Connelly y Clandinin, 1995; Mateos y Núñez, 2011).

Este enfoque ha demostrado gran valor en el ámbito educativo, particularmente en la EF, donde la narrativa permite visibilizar aspectos emocionales, sociales y corporales frecuentemente invisibilizados por métodos más tradicionales. Estudios recientes han utilizado relatos, poemas, cuentos o incluso canciones como formas válidas de expresión investigativa, demostrando la potencia de este enfoque para capturar la complejidad de las experiencias escolares de manera rica y significativa (Carless y Douglas, 2011; Devís-Devís, 2017; Dowling et al., 2013; Fitzpatrick, 2018; Martos-García et al., 2023).

3.2.2. Técnicas e instrumentos de recolección de datos etapa 1

3.2.2.1. Relato narrativo

Para abordar el objetivo de caracterizar las experiencias vividas sobre evaluación en las clases de EF escolar por estudiantes de una Facultad de Educación, se empleó como técnica de recolección de información el relato narrativo. Esta técnica se inscribe en la tradición de la investigación narrativa, que reconoce el poder del relato como vía privilegiada para explorar el sentido que las personas otorgan a sus vivencias. En este caso, se solicitó al estudiantado que escribiera microrrelatos donde describieran experiencias significativas de evaluación en Educación Física, permitiendo acceder a sus emociones, creencias, prácticas y a la reconstrucción de contextos escolares (Bolívar, 2002; Connelly y Clandinin, 1990; León, 2015). El uso del microrrelato como instrumento resulta especialmente pertinente en investigaciones educativas, ya que condensa en un formato breve pero profundo una experiencia concreta, estimulando una expresión subjetiva y situada. Como lo destacan Roca y Pardo (2023), este tipo de formato narrativo, al tiempo que favorece una comunicación accesible y atractiva, también promueve una participación proactiva del emisor y del lector, convirtiéndose en una herramienta valiosa para recoger experiencias educativas con intensidad expresiva y sentido personal.

En este estudio, el instrumento fue un formulario digital, que entregaba un apartado para la caracterización del estudiantado participante, y luego una consigna abierta que invitó a los y las participantes a relatar, en primera persona, una vivencia de evaluación que consideraran significativa durante su etapa escolar. Cabe destacar que la aplicación del instrumento digital se aplicó de forma presencial en una sala de computación que permitía al estudiantado expresarse en un espacio de tranquilidad y confianza.

3.2.2.2. Grupo focal

Para abordar el objetivo de indagar sobre la influencia del proceso evaluativo de la EF escolar en las actitudes del alumnado hacia la asignatura, se empleó la técnica cualitativa de grupo focal, la cual permite explorar en profundidad percepciones, significados y construcciones sociales a través de la interacción entre los y las participantes. Esta técnica consiste en reunir a un grupo pequeño de personas para discutir, con la guía de un moderador,

un conjunto de preguntas centradas en un tema específico, generando un espacio donde las opiniones de unos estimulan las respuestas de otros, permitiendo así una producción colectiva de significados (Thofehrn et al., 2013).

En el contexto de la investigación educativa, el grupo focal es altamente pertinente, ya que posibilita el acceso a los discursos, emociones y vivencias de los sujetos en torno a fenómenos complejos, como lo es la evaluación escolar, revelando matices que muchas veces no emergen en entrevistas individuales o encuestas (Arias, 2022; Barredo et al., 2023).

En investigaciones con enfoque narrativo, los grupos focales complementan otras técnicas como los relatos escritos, al permitir contrastar, profundizar o ampliar las experiencias previas relatadas, y dar lugar a una construcción dialógica del sentido (Porta y Flores, 2020).

En esta tesis doctoral, el guion del grupo focal fue construido a partir del análisis inductivo de los relatos narrativos previos, permitiendo así que las temáticas emergentes del estudiantado guiaran la conversación grupal. Posteriormente, este guion fue validado mediante juicio de expertos, quienes evaluaron su claridad, coherencia y capacidad para fomentar la participación activa del grupo. Esta estrategia metodológica garantizó una conexión coherente entre técnicas e instrumentos, fortaleciendo la profundidad del análisis y la credibilidad de los hallazgos.

3.2.3. Participantes etapa 1

Dado el enfoque de los objetivos de investigación propuestos para la etapa 1 de esta tesis doctoral, se decidió convocar a personas que hubieran cursado recientemente la totalidad de su etapa escolar, con el fin de asegurar que su experiencia en la asignatura de EF fuese completa y actual. Para ello, se seleccionó específicamente a estudiantes de una universidad chilena, específicamente en la región del Maule, que cursaban el quinto semestre de sus respectivas carreras, ya que su formación escolar no se vio interrumpida por las cuarentenas ni por la modalidad virtual derivada de la pandemia por COVID-19, considerando que su ingreso a la universidad coincidió con el inicio de esta crisis sanitaria.

La Facultad de Educación de la universidad participante está compuesta por seis especialidades pedagógicas, entre ellas la carrera de Pedagogía en EF. Con el propósito de evitar una sobrerrepresentación de estudiantes de dicha especialidad (lo que podría generar sesgos en las percepciones recogidas debido a su afinidad disciplinar), se optó por incluir participantes de todas las carreras de Pedagogía. Esta decisión buscó asegurar una representación equilibrada tanto en términos de especialidad como de género, promoviendo así una mayor diversidad de miradas y experiencias que enriquecieran el análisis y comprensión del fenómeno investigado. En investigaciones cualitativas, este tipo de muestreo intencionado y equilibrado es especialmente valorado, ya que permite capturar la riqueza y heterogeneidad de los significados atribuidos por los y las participantes, atendiendo a la variedad de contextos, trayectorias y subjetividades (Maxwell, 2013).

Finalmente respondieron al llamado voluntario 60 estudiantes de la Facultad de Educación antes mencionada, pertenecientes a 6 grados diferentes: Pedagogía en Lengua Castellana y Comunicación, Pedagogía en inglés (PI), Pedagogía en Educación Física (PEF), Pedagogía en Educación Parvularia (PEP), Pedagogía General Básica (PGB) y Pedagogía en Historia y Geografía (PHG). En el instrumento se consultó por el género, dando las opciones de femenino, masculino y no binario, resultando la siguiente distribución presentada en la tabla 6.

Tabla 6. Distribución según género de las personas participantes.

Disciplina	Femenino	Masculino	Cantidad
Pedagogía en Lengua Castellana y Comunicación	7	3	10
Pedagogía en Educación Física	3	7	10
Pedagogía en Inglés	5	5	10
Pedagogía en Educación Básica	7	3	10
Pedagogía en Historia y Geografía	3	7	10
Pedagogía en Educación Parvularia	10	0	10
Total	35 (58%)	25 (42%)	60 (100%)

En base a su experiencia y formación en pedagogía, en el año 2022 se invitó a participar al alumnado del nivel 5 de cada grado, lo que corresponde con el primer semestre del tercer

año de su formación, en donde según el plan curricular, todos y todas deberían tener formación en la asignatura de didáctica y evaluación. Por tanto, la edad de los y las participantes oscilaba entre 20 y 30 años.

Dentro de los criterios de inclusión, se tuvo en cuenta que el estudiantado hubiera vivenciado la asignatura de EF en su edad escolar de forma regular, es decir, sin haberse eximido de ello durante un periodo mayor a un año.

Es relevante destacar que el estudiantado de PEF constituye solo un sexto del total de participantes. Esta distribución fue intencionada por el equipo de investigación para evitar sesgar los resultados hacia las experiencias y percepciones asociadas únicamente a la carrera de PEF, buscando obtener una perspectiva más amplia y diversa.

3.2.3.1. Participantes grupos focales

Para la segunda técnica de recolección de datos de esta etapa, se realizaron dos grupos focales con un total de 18 estudiantes participantes. La convocatoria se dirigió inicialmente a los 60 estudiantes que habían respondido previamente a la consigna de los relatos narrativos; sin embargo, solo 18 de ellos aceptaron participar en los grupos focales. Esta muestra estuvo conformada por estudiantes de diferentes disciplinas pedagógicas, distribuidos proporcionalmente como se observa en la Figura 1. En términos de género, participaron 11 mujeres (61%) y 7 hombres (39%), tal como se detalla en la Tabla 7.

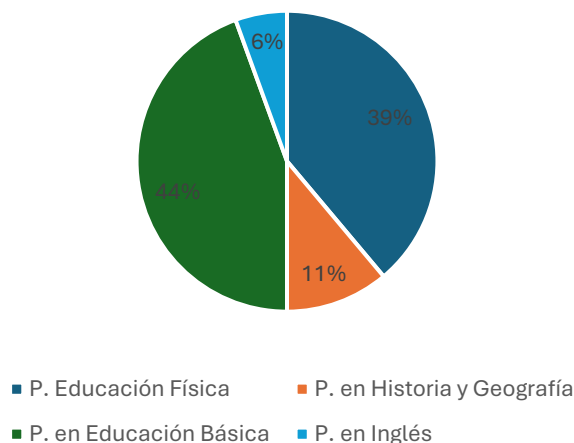
Tabla 7. Participación Grupos focales.

Grupo focal	Mujeres	Hombres	Total
1	3	5	8
2	8	2	10
Total	11 (61%)	7 (39%)	18 (100%)

A nivel disciplinar, el grupo estuvo compuesto por estudiantes de Pedagogía en Educación Básica (44%), Pedagogía en Educación Física (39%), Pedagogía en Historia y Geografía (11%) y Pedagogía en Inglés (6%). Esta distribución permitió recoger una diversidad de perspectivas en relación con sus experiencias evaluativas en la asignatura de EF, aportando riqueza al análisis cualitativo posterior.

Figura 2. Distribución por disciplina en los grupos focales.

Participación Grupos Focales



3.2.4. Análisis de datos etapa 1

El análisis de datos de la etapa 1 parte con los datos obtenidos en el primer instrumento de relatos narrativos escritos. Cada uno de los 60 participantes entregó sus datos de caracterización y luego procedió a redactar sus experiencias que recordaban en evaluación en la asignatura de EF en edad escolar.

Un primer acercamiento para la caracterización de la muestra utilizó porcentajes y cantidad de participantes según edad, género, especialidad en sus estudios de pedagogía, etc. Todo este análisis previo para contextualizar a las personas participantes se realizó con el software Microsoft Excel versión 2502 (compilación 18526.20168), correspondiente a la suite de Microsoft 365 para empresas. Los gráficos y tablas obtenidas se pueden apreciar en el apartado de participantes de la etapa 1.

Para la organización preliminar de los 60 relatos narrativos, se creó un documento en el software Microsoft Word versión 2502, lo cual permitió en primera instancia corregir los errores ortográficos o gramaticales claros, manteniendo fielmente el relato del estudiantado. Posterior a ello se realizó una lectura exploratoria y general de todos los relatos, con el objetivo de familiarizarse con el contenido y captar la diversidad de experiencias y tonos

abierta de los relatos, en la cual se identificaron unidades de significado relevantes (Boeije, 2010). Estas unidades correspondían a fragmentos textuales que expresaban emociones, juicios, contextos o acciones relacionadas con sus vivencias en evaluación en EF en edad escolar. Es un buen momento para señalar que, al permitir expresarse de una forma tan abierta y libre, se encontraron relatos o vivencias que, si bien eran contextualizadas a la EF, no tenían concordancia específica con la evaluación, pero aun así sirvieron como contextualización del relato.

Cada unidad textual identificada del documento de Word, fue etiquetada con un código que describía de manera breve su contenido. En el caso de esta investigación se optó por añadir los códigos al costado del relato, asignado un color diferenciador a cada fragmento textual, tal como se puede apreciar en la figura 3.

Una vez realizada esta codificación inicial, los códigos fueron agrupados en categorías preliminares a través de un proceso de comparación constante (Corbin Y Strauss, 2015). Esto implicó examinar las similitudes y diferencias entre los relatos para establecer conexiones conceptuales. Este procedimiento permitió construir una estructura temática que articuló las experiencias en torno a núcleos interpretativos, como el tipo de evaluación recibida, las emociones asociadas, el rol del profesorado, y las consecuencias pedagógicas y personales.

Una vez definidas estas categorías preliminares, hubo una segunda instancia de organización de las unidades textuales identificadas en la codificación abierta; esta vez serían organizadas por contenido curricular, identificando si las experiencias matizaban percepciones negativas o desagradables, positivas o agradables o bien era neutras, lo que generó la posibilidad de agrupar las vivencias y categorizarlas de una forma ordenada.

A partir de esta nueva codificación, se elaboraron tanto un mapa de relaciones temáticas (figura 4), que considera la totalidad de categorías emergentes de los relatos narrativos, y una matriz en la que se organizaron las categorías y subcategorías emergentes relacionadas a la evaluación de la EF según los contenidos curriculares, permitiendo visualizar de manera estructurada la frecuencia de aparición de los distintos códigos identificados en los relatos del estudiantado. Esta representación facilitó tanto la sistematización de los hallazgos como la interpretación comparativa entre experiencias. Tal como se presenta en la tabla 8, el uso

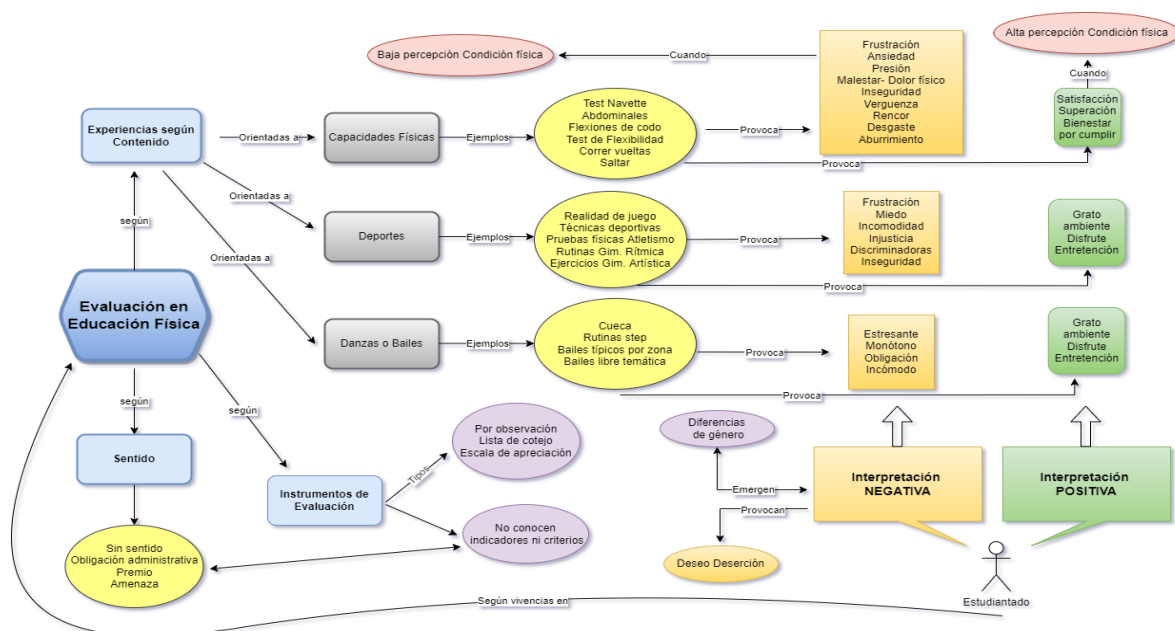
de matrices es una estrategia recomendada para analizar y visualizar datos cualitativos de forma ordenada, ya que permite observar patrones, relaciones y jerarquías entre los códigos generados (Miles et al., 2014).

Tabla 8. Matriz de categorías emergentes de los relatos narrativos, desde los contenidos curriculares.

Categoría		Ejemplos mencionados		Percepciones negativas		Percepciones positivas	
Condición (60)	Física	Test Course Navette	(27)	Frustración	(10)	Satisfacción	(10)
		Abdominales	(14)	Desmotivación	(8)	Superación	(5)
		Correr	(14)	Enojo	(6)		
		Flexibilidad	(5)	Vergüenza	(6)		
		Salto de Cuerda	(3)	Malestar físico	(6)		
		Test de Cooper	(3)	Inseguridad	(5)		
		Salto de Cooper	(2)	Desgaste físico	(3)		
		Burpees	(2)	Inferioridad	(3)		
		Circuitos	(2)				
Deportes (28)		Deporte genérico	(12)	Frustración	(8)	Satisfacción	(6)
		Gimnasia	(10)	Discriminación género	(4)	Compañerismo	(4)
		Fútbol	(8)	Desmotivación	(3)	Diversión	(3)
		Voleibol	(7)	Inseguridad	(3)	Logro	(3)
		Básquetbol	(7)	Discriminación general	(3)		
		Handball	(5)	Incomodidad	(3)		
		Atletismo	(2)	Miedo	(2)		
		Natación	(2)	Ansiedad	(2)		
				Vergüenza	(2)		
				Injusticia	(2)		
Danza /Baile (24)		Cueca	(10)	Obligación	(6)	Entretención	(5)
		Bailes tradicionales	(10)	Estresante	(4)	Satisfacción	(3)
		Coreografías variadas	(3)	Incomodidad	(4)		
		Danzas de Rapa Nui	(2)	Monotonía	(3)		

Elaboración propia en base a los relatos narrativos. El número () corresponde a la frecuencia de concurrencia de los 60 relatos.

Figura 4. Interpretaciones estudiantiles sobre la evaluación en EF: experiencias, sentidos y consecuencias.



Elaboración propia en base a los relatos narrativos.

Tal como se indicó en el apartado de técnicas e instrumentos de la etapa 1, las categorías emergidas del análisis de los relatos narrativos sirvieron como base para la elaboración del guion de los grupos focales realizados con el estudiantado participante. Las transcripciones de dichos grupos fueron analizadas mediante un procedimiento similar al aplicado en los relatos escritos, siguiendo un enfoque inductivo y temáticamente orientado. En primer lugar, se llevó a cabo una lectura holística del material con el fin de captar el tono general y las experiencias dominantes (Miles et al., 2014). Posteriormente, se realizó una codificación abierta que permitió identificar unidades de significado relevantes según los temas abordados en la discusión grupal. Finalmente, los códigos fueron organizados en una matriz de categorías y subcategorías, integrando tanto los elementos emergentes como aquellos ya identificados en el primer análisis, lo que favoreció la triangulación y la consolidación de patrones interpretativos. La matriz presentada en la tabla 9, da cuenta de las categorías y subcategorías que emergieron de dicho análisis.

Tabla 9. Matriz de categorías emergentes desde los Grupos Focales.

Categorías	Subcategorías	Códigos emergentes
Metodología	Estilos de clase	Enfoque deportivo y realidad de juego Acondicionamiento físico Tiempo libre en hora de clases En edad infantil, gran relevancia al juego
	Planificación	Escasa claridad o ausencia de los objetivos Cambios de temas de forma brusca Falta de continuidad en las unidades
Evaluación	Propósitos	Solo para calificar Como recompensa por competir o participar Sin explicar el sentido Por cumplir Por medir la condición física Fines formativos Retroalimentación en base al disfrute de la clase
	Instrumentos	Rúbricas Observación (sin registro) Lista de cotejo Tablas de rendimiento físico Completa ausencia de instrumentos o indicadores
Percepciones	Injusticia social	Comentarios de burla Estereotipos de género en deportes Comentarios de discriminación por sobrepeso u obesidad Sentimiento de discriminación por competencia motriz
	Deseo de deserción	Por sentirse torpe Por injusticia Por incomodidad (por no rendir) Monotonía Escasa libertad Excusas con apoyo de padres

Elaboración propia en base a los grupos focales

3.3. Metodología etapa 2

3.3.1. Enfoque de investigación

La segunda etapa de esta tesis doctoral aborda los objetivos específicos 3 y 4, orientados a comprender las conceptualizaciones, perspectivas y prácticas evaluativas del profesorado de EFyS en ejercicio, tanto en el ámbito escolar como universitario. Para este propósito, se adopta un enfoque cualitativo de corte interpretativo, el cual privilegia el análisis profundo de las experiencias vividas, los significados atribuidos y las dinámicas que configuran el hacer docente en contextos educativos reales (Creswell y Poth, 2018; Taylor y Bogdan, 1986).

Este enfoque parte de la premisa de que la realidad educativa es construida socialmente y que los conocimientos relevantes para su comprensión emergen de la interacción entre las y los sujetos, sus contextos y sus discursos (Denzin y Lincoln, 2011; Hernández Carrera, 2014). Así, la interpretación de las percepciones del profesorado permite desvelar tanto las intenciones que guían sus acciones evaluativas como las tensiones, contradicciones y sentidos que les atribuyen en la práctica cotidiana. Tal como señalan Creswell y Poth (2018), la investigación cualitativa de tipo interpretativo resulta especialmente pertinente cuando se busca comprender fenómenos educativos desde la perspectiva de quienes los viven, reconociendo el valor de sus narrativas para construir conocimiento situado y significativo.

Este enfoque reconoce al profesorado como agente epistémico, es decir, como conocedor legítimo de su práctica y productor de saberes pedagógicos, cuyas voces resultan indispensables para comprender los procesos evaluativos en EF en su complejidad (Cornejo y Salas, 2011; Monforte y Pérez-Samaniego, 2021). De esta forma, la investigación interpretativa no se limita a describir hechos, sino que busca interpretar las estructuras de significación que los sustentan, atendiendo a la subjetividad, la reflexividad y el contexto desde el cual emergen (Maxwell, 2013).

3.3.2. Técnicas e instrumentos de recolección de datos etapa 2

3.3.2.1. Entrevista

La técnica de la entrevista es ampliamente utilizada en la investigación cualitativa, ya que permite acceder al mundo subjetivo de los y las participantes, explorando sus pensamientos, emociones y experiencias desde sus propios marcos de referencia (Rubin y Rubin, 2020). Esta técnica puede adoptar distintas modalidades según el grado de estructuración: entrevistas estructuradas, semiestructuradas y abiertas (Evans, 2018). Las entrevistas estructuradas siguen un guion rígido, mientras que las abiertas ofrecen mayor libertad de expresión. Las entrevistas semiestructuradas, en cambio, equilibran la guía temática con la flexibilidad para profundizar en temas emergentes, lo que las hace especialmente adecuadas para investigaciones educativas centradas en las experiencias y significados construidos por los actores involucrados (Troncoso y Daniele, 2003).

En el marco de esta investigación, se optó por entrevistas semiestructuradas debido a su capacidad para explorar en profundidad las concepciones y experiencias evaluativas del profesorado de EF, en coherencia con los objetivos 3 y 4 del estudio. Este tipo de entrevistas ha sido ampliamente utilizado en investigaciones internacionales sobre evaluación en EF, permitiendo captar las racionalidades docentes desde sus propios relatos. Estudios como los de Hay y Penney (2009), Alfrey y Gard (2014) y Aarskog (2020) han empleado esta técnica para indagar en tensiones pedagógicas, participación del estudiantado y prácticas evaluativas contextualizadas. Además, la revisión de Killian y Woods (2021) destaca su uso como estrategia metodológica central en el campo.

El diseño del guion de la entrevista se construyó siguiendo un enfoque sistemático y fundamentado (Troncoso y Daniele, 2003). En un primer momento, se realizó una revisión exhaustiva de la literatura especializada en evaluación en el ámbito de la Educación Física. Este análisis permitió identificar los principales conceptos, enfoques y desafíos relacionados con la temática, sirviendo como base para estructurar las preguntas del guion. Posteriormente, se esbozó un esquema de preguntas con el propósito de responder al objetivo principal de la investigación. El guion se organizó en cuatro secciones principales: (1) una introducción orientada a la caracterización del contexto de las y los entrevistados, (2) una exploración sobre su conceptualización de la evaluación, (3) un análisis de sus experiencias evaluativas

en la etapa escolar y durante la formación universitaria, y (4) una indagación sobre sus prácticas evaluativas actuales en el ejercicio profesional.

Una vez diseñado, el guion fue sometido a un proceso de análisis de calidad mediante juicio de expertos. Para ello, se consultó a especialistas en el área, quienes evaluaron la pertinencia, claridad y capacidad de las preguntas para captar información relevante y consistente con los objetivos de la investigación. Las observaciones y recomendaciones de los expertos fueron incorporadas para garantizar la calidad y adecuación del instrumento. Finalmente, la entrevista quedó confeccionada para que su duración rondara los 45 minutos en la aplicación.

3.3.2.2. Análisis documental

En el marco de una investigación cualitativa de enfoque interpretativo, el análisis documental se constituye como una técnica rigurosa y complementaria que permite profundizar en la comprensión de los significados construidos por los sujetos, mediante la revisión de documentos generados en contextos reales y educativos. Esta técnica se sustenta en un proceso sistemático, crítico y reflexivo de análisis de materiales escritos que contienen información relevante para el objeto de estudio, tales como registros institucionales, textos académicos, documentos oficiales o instrumentos pedagógicos (Finol de Franco y Arrieta, 2021; Sánchez y Vega, 2003). Su valor metodológico reside en la posibilidad de triangular la información obtenida por medio de otras técnicas, como las entrevistas en profundidad, con evidencias tangibles que reflejan prácticas, discursos y decisiones pedagógicas de los actores involucrados. En el ámbito educativo, el análisis documental no solo permite identificar estructuras, patrones y sentidos presentes en los materiales revisados, sino también contextualizar las acciones docentes y explorar cómo se expresan sus concepciones pedagógicas y evaluativas a través de dichos instrumentos (Finol de Franco y Arrieta, 2021).

En la segunda etapa de la tesis doctoral, el análisis documental se focalizó en los instrumentos de evaluación facilitados voluntariamente por parte del profesorado participante, los cuales habían sido utilizados durante el año académico en curso. Esta estrategia metodológica permitió complementar y contrastar los relatos obtenidos en las entrevistas, aportando una visión más integrada y situada de las conceptualizaciones y

prácticas evaluativas en EF. Tal como señalan Bermeo-Yaffar et al. (2016), el análisis documental no solo contribuye con evidencia empírica relevante, sino que también fortalece el desarrollo de competencias investigativas en el ámbito docente, consolidando así la calidad y solidez del proceso investigativo.

Además del análisis de los instrumentos de evaluación utilizados por el profesorado, en el Estudio 4 se incorporó un análisis documental de naturaleza cuantitativa, centrado en las calificaciones oficiales en la asignatura de EFyS. Se analizaron registros correspondientes a estudiantes desde primero básico hasta cuarto medio (6 a 18 años), proporcionados por el MINEDUC de Chile mediante una solicitud de acceso a la información pública.

El periodo analizado abarcó desde 2018 a 2022, delimitación que responde tanto a criterios normativos como metodológicos. El año 2018 marca la entrada en vigencia del Decreto Exento N.º 67, el cual redefine los principios que orientan la evaluación, calificación y promoción escolar en el país. Por su parte, el año 2022 corresponde al momento en que se realizó el trabajo de campo con el profesorado, lo que permite establecer un marco temporal coherente con el resto del estudio.

Las bases de datos utilizadas contienen información sobre los promedios anuales de calificaciones por curso, nivel educativo y región, incluyendo la totalidad de los subsectores del currículum, con énfasis específico en EFyS. La información expuesta en la tabla 10 permitió examinar tendencias y patrones en la distribución de notas, tanto a nivel nacional como regional, con el objetivo de contrastar las percepciones docentes con evidencia empírica.

Tabla 10. Bases de datos de calificaciones utilizadas en este estudio.

Año escolar	Número total de registros	Número de establecimientos
2018	30612736	9160
2019	32227180	9068
2020	25196613	9040
2021	32410770	8964
2022	32793724	8903

Elaboración propia

3.3.3. Participantes etapa 2

Los participantes fueron seleccionados mediante un muestreo intencional (Palinkas et al., 2015), el cual permite elegir sujetos que tienen una experiencia significativa y pertinente para el fenómeno de interés, asegurando así la relevancia de los datos recogidos. Los participantes fueron 3 profesoras y 7 profesores de EFyS que, en el momento de la entrevista, se encontraban ejerciendo en diferentes establecimientos educativos de la ciudad de Talca. Dentro de los criterios de inclusión se consideró, como principal requisito, llevar al menos 1 año ejerciendo como docente de la asignatura mencionada en cualquier contexto escolar.

Tabla 11. Caracterización del profesorado entrevistado.

Participante	Género	Rango edad (años)	Años de experiencia docente	Dependencia del centro educativo
1	Mujer	50 a 59	32	Particular
2	Mujer	40 a 49	24	Municipal
3*	Hombre	30 a 39	11	Particular subvencionado
4	Hombre	30 a 39	10	Municipal
5	Hombre	30 a 39	10	Particular
6*	Hombre	30 a 39	8	Particular subvencionado
7*	Hombre	30 a 39	7	Particular subvencionado
8	Hombre	30 a 39	4	Municipal
9	Hombre	20 a 29	4	Municipal
10*	Mujer	20 a 29	2	Municipal

* Envío instrumento de evaluación.

La Tabla 11 presenta una caracterización del profesorado participante, señalando quienes de ellos compartieron sus instrumentos de evaluación para el análisis. Estos instrumentos incluyen una rúbrica, dos escalas de apreciación y un registro diario, dirigidos a los cursos de 1° y 2° básico. Los objetivos varían entre evaluar habilidades motrices básicas (como la manipulación y la combinación de locomoción y manipulación en circuitos), y hábitos de higiene, conducta, participación y esfuerzo. La cantidad de indicadores para cada instrumento oscila entre 5 y 6, evidenciando un enfoque detallado en la evaluación de aspectos específicos relacionados con las habilidades y actitudes de los y las estudiantes.

3.3.4. Análisis de datos etapa 2

3.3.4.1. Análisis cualitativo

El análisis de datos correspondiente a la segunda etapa de la investigación doctoral se centró en la información obtenida a partir de entrevistas semiestructuradas aplicadas a 10 docentes de EF en ejercicio, junto con el análisis documental de los instrumentos de evaluación que algunos de ellos facilitaron. A diferencia del análisis más inductivo realizado en la etapa 1, en esta fase el procedimiento analítico estuvo guiado por las temáticas estructuradas previamente en el guion de entrevista, lo que implicó un abordaje más focalizado y temáticamente orientado desde el inicio del análisis.

Las entrevistas, de carácter individual y presencial, fueron grabadas con autorización de los participantes y posteriormente transcritas íntegramente de forma manual en documentos de texto. El proceso de análisis se realizó sin el uso de software especializado, siguiendo un enfoque sistemático y artesanal, lo que permitió un contacto más directo y reflexivo con el material empírico. La lectura inicial del corpus se realizó de forma holística (Miles, Huberman y Saldaña, 2014), con el fin de identificar la riqueza contextual de cada relato y establecer relaciones entre los discursos de los y las docentes.

A continuación, se procedió a una codificación temática, organizada en torno a las cuatro secciones principales del guion de entrevista: (1) caracterización del contexto docente, (2) conceptualización de la evaluación, (3) experiencias personales de evaluación durante su etapa escolar y universitaria, y (4) prácticas evaluativas actuales. Esta estructura orientó la segmentación del contenido, permitiendo identificar unidades de significado en cada eje temático.

En cada segmento, se aplicó una codificación abierta manual, desde la cual emergieron subtemas y regularidades interpretativas, dando lugar a categorías y subcategorías analíticas. Estas fueron registradas en matrices de análisis diseñadas para cada una de las secciones, que incluyeron citas textuales representativas y observaciones interpretativas por parte del equipo investigador. Este trabajo fue acompañado de un proceso de triangulación por investigador, contrastando las interpretaciones y consensuando las categorías finales, con el objetivo de fortalecer la credibilidad del análisis (Smith y McGannon, 2018).

Para organizar la categorización, se optó por la elaboración de matrices por temática abordada para, posteriormente, integrar dichas matrices junto al análisis documental posterior, con la finalidad de evitar la segregación de resultados según el instrumento. La matriz de categorías de la tabla 12 corresponde al análisis de las experiencias escolares, universitarias y de su propia práctica como docente.

Tabla 12. Matriz de categorías sobre experiencias del profesorado en evaluación.

Categoría	Subcategoría	Códigos	Descripción
Experiencia Escolar	Evaluación Tradicional (9)	Test físico Técnicas deportivas Rendimiento	En edad escolar el profesorado fue evaluado principalmente por criterios estrictos de rendimiento y técnica.
	Enfoque Actitudinal (Ausente) (5)	Solo desempeño Rendimiento	Falta de reconocimiento del esfuerzo o la actitud, se centraban únicamente en los resultados cuantificables.
Experiencia Universitaria	Evaluación Tradicional (8)	Técnicas deportivas Rendimiento Pruebas tipo test	En la universidad, la evaluación se enfocaba mucho en la técnica y en el rendimiento en deportes específicos.
	Reflexión y Cambio (6)	Reflexión Asignaturas clave Profesorado reflexivo	La universidad fue un punto de inflexión para algunos profesores, donde empezaron a cuestionar la evaluación tradicional.
	Enfoque Actitudinal emergente(4)	Considerar actitud Valor por esfuerzo Enfoques alternativos	Algunos cursos en la universidad mostraron la importancia de evaluar no solo el rendimiento, sino también el proceso.
	Cambio en la Práctica Docente (10)	Realidad Respuesta del estudiantado Observación y reflexión	A partir de su práctica docente, muchos profesores cambiaron su enfoque de evaluación hacia uno más inclusivo y adaptativo.
Experiencia como Profesor(a)	Evaluación Tradicional (Limitada) (3)	Rendimiento Rendimiento + esfuerzo Rendimiento +actitud	A pesar del cambio en su enfoque, algunos profesores siguen empleando criterios tradicionales en ciertos contextos.
	Enfoque Actitudinal e Inclusivo (7)	Esfuerzo Actitud Participación	Los profesores enfatizan la importancia de la actitud, el esfuerzo, y el proceso de aprendizaje por encima del resultado final.

Elaboración propia desde el análisis a las entrevistas del profesorado.

De forma separada, pero en concordancia con la matriz anterior, se generó la matriz de la tabla 13, que corresponde al análisis respecto a las respuestas asociadas a la conceptualización y prácticas evaluativas.

Tabla 13. Matriz de categorías Conceptualización y prácticas evaluativas.

Categorías	Subcategorías	Códigos	Descripción
Conceptualización	Formativa	Progreso Proceso Comunicada Formativa Más que calificar	Se describe la evaluación como un proceso, que debe considerar un seguimiento en un periodo de tiempo, para comunicar los avances.
		Individual Personalizada Intereses estudiante Contexto Adecuada Accesible	Se enfatiza en adecuar y considerar las características individuales del estudiantado en la evaluación, accesible para todos los niveles de competencia motriz.
	Foco	Cualitativo Esfuerzo Información Transición Conceptual Procedimental Actitudinal	La concepción de evaluación se focaliza principalmente en la dimensión actitudinal, considerando el esfuerzo y la participación del estudiantado. En menor medida se mencionan las dimensiones procedimentales y conceptuales.
	Frecuencia	Una por Unidad 2 a 3 al semestre Siempre	Se menciona que la evaluación está presente en todo momento, pero la calificación se genera 2 a 3 veces por semestre. Es decir, de 4 a 6 calificaciones anuales.
	Instrumentos	Registro clase a clase Escala de apreciación Lista de cotejo Rúbricas Validados por UTP	Se mencionaron los registros clase a clase por observación, escalas de apreciación, rúbrica y listas de cotejo. Siempre pasando por un proceso de validación interno por UTP.
		Actitudinal Comportamiento Participación Conceptuales Esfuerzo Superación Procedimentales	Se menciona con mayor frecuencia componentes de la dimensión actitudinal como elemento clave y de gran ponderación en la calificación. Se mencionan además las dimensiones procedimentales y conceptuales en menor medida.
Prácticas evaluativas	Usos	Diagnóstico Formativo Calificar Seguimiento de actitud Observación clase a clase	Los usos atribuidos a la evaluación van desde el diagnóstico, seguimiento clase a clase, evaluaciones de procesos y calificación.

Elaboración propia en base a las entrevistas del profesorado EFyS

En paralelo, se realizó un análisis documental de los instrumentos de evaluación proporcionados por los cuatro docentes mencionados anteriormente, se presenta la caracterización de dichos instrumentos en la tabla 14.

Tabla 14. Caracterización de los instrumentos de evaluación analizados.

Participante (N de tabla 1)	Tipo de instrumento	Nivel educativo	Objetivo del instrumento	Cantidad de indicadores
3	Rúbrica	1° básico (primaria)	Evaluar habilidades motrices básicas, especialmente manipulación, a través de circuitos.	6
6	Escala apreciación	2°básico (primaria)	Evaluar habilidades motrices de locomoción y manipulación, promoviendo participación activa.	6
7	Escala apreciación	2°básico (primaria)	Evaluar el combinar habilidades motrices de locomoción y manipulación en circuitos, con buena actitud.	5
10	Registro diario	1°básico (primaria)	Evaluar hábitos de higiene, conducta, participación y esfuerzo,	4

Elaboración propia

Estos documentos fueron examinados en función de los criterios e indicadores de evaluación detallados en los mismos, con especial atención a las dimensiones actitudinal, procedimental y conceptual. Desde este análisis subyacen categorías en base a los códigos que se describen en la tabla 15, sintetizando la información en una matriz de categorías.

Tabla 15. Matriz de categorías desde los instrumentos de evaluación.

Categoría	Código	Descripción	Ejemplo del Instrumento
Dimensión actitudinal	Participación activa	Evaluación de la disposición del estudiante para participar activamente en las actividades de clase.	"Participa activamente en las actividades de la clase poniendo atención y respetando turnos" (Rúbrica 1° Básico)(Escala ap. 2° Básico).
	Seguimiento de instrucciones	Capacidad del estudiante para escuchar y seguir las instrucciones del docente.	"El estudiante escucha y sigue las instrucciones entregadas por el profesor manifestando dudas de manera clara" (Pauta de evaluación 2°)(Registro diario).
	Entusiasmo y compromiso	Evaluación del entusiasmo y el compromiso mostrado por el estudiante durante las actividades.	"Demuestra entusiasmo y compromiso en todas las actividades" (Rúbrica 1° Básico) (Registro diario).
	Trabajo en equipo	Evaluación del trabajo conjunto y respeto de turnos y opiniones de otros estudiantes.	"El estudiante participa en conjunto con su grupo manteniendo un orden en los turnos" (Escala ap. 2°).
Dimensión procedimental	Control de objetos	Habilidad del estudiante para manipular y controlar objetos en el circuito.	"Ejecución de circuito demostrando control en los diferentes objetos" (Escala ap, Rúbrica)
	Habilidad motriz combinada	Combinación de habilidades de locomoción y manipulación, evaluando fluidez y coordinación.	"Corre en el circuito de manera fluida lanzando y atrapando el implemento con una o ambas manos" (Rúbrica 1° Básico)
Dimensión conceptual	Reconocimiento de habilidades	Reconocimiento de las habilidades motrices utilizadas en el circuito (lanzar, atrapar, etc.).	"El estudiante reconoce con facilidad la habilidad motriz trabajada" (Pauta de evaluación 2°).

Durante todo el proceso analítico se sostuvo un registro reflexivo sistemático de las decisiones metodológicas, lo que permitió documentar la evolución interpretativa y explicitar la reflexividad del equipo investigador como parte integral del enfoque cualitativo (Finlay, 2002; Olmos-Vega et al., 2022). Esta práctica favoreció la trazabilidad del análisis y la transparencia epistemológica, fortaleciendo la coherencia entre los hallazgos y las perspectivas teóricas que los sustentan.

El análisis de esta etapa posibilitó una reconstrucción profunda de las relaciones entre las experiencias previas de evaluación vividas por el profesorado, sus concepciones actuales y las prácticas concretas que despliegan en el aula. Este proceso reveló continuidades, tensiones y transformaciones en torno al acto de evaluar en EF, evidenciando la complejidad de sus posicionamientos pedagógicos.

Asimismo, el contraste entre los discursos docentes y los documentos institucionales permitió triangular las prácticas declaradas con aquellas efectivamente evidenciadas en sus instrumentos de evaluación, enriqueciendo la interpretación mediante una lectura situada y crítica de los datos (Bowen, 2009). A partir de este cruce analítico, se elaboró una matriz integrada que articula los códigos emergentes provenientes de las entrevistas y los instrumentos revisados, tal como se presenta en la Tabla 16, donde se organizan las dimensiones, categorías, códigos y sus respectivas descripciones.:

MARCO METODOLÓGICO

Tabla 16. Matriz de categorías y subcategorías emergentes del análisis de las entrevistas e instrumentos de evaluación.

Dimensión	Categoría	Subcategoría	Códigos	Descripción
Experiencias previas Evaluación	Experiencias en edad escolar	Evaluación tradicional	Criterios estrictos (Ent) Rendimiento físico (Ent) Técnica deportiva (Ent)	En edad escolar, el profesorado fue evaluado principalmente por criterios estrictos de rendimiento y técnica.
		Ausencia evaluación	No reconoce esfuerzo (Ent) Solo procedimental (Ent) No se evaluaba (Ent)	Falta de reconocimiento del esfuerzo o la actitud, se centraban únicamente en los resultados cuantificables.
		Evaluación tradicional	Rendimiento técnico (Ent) Asignaturas enfocadas en rendimiento (Ent)	En la universidad, la evaluación se enfocaba mucho en la técnica y en el rendimiento en deportes específicos.
	Experiencias en universitarias	Reflexión en enfoque	Cuestionamiento a métodos tradicionales (Ent) Consideración a la diversidad (Ent)	La universidad fue un punto de inflexión para algunos profesores, donde empezaron a cuestionar la evaluación tradicional.
		Enfoque actitudinal emergente	Enfoque en proceso (Ent) Asignaturas clave en enfoque actitudinal (Ent)	Algunos cursos en la universidad mostraron la importancia de evaluar no solo el rendimiento, sino también el proceso.
	Transición de enfoque	Cambios en sus propuestas	De la ausencia a evaluar (Ent) De enfoque tradicional a integral (Ent)	La totalidad del profesorado entrevistado manifiesta algún nivel de cambios entre lo que vivenció en sus etapas escolar y/o universitaria.
	Proceso evaluativo	Conceptualización	Progreso (Ent+Ins) Proceso (Ent+Ins) Comunicada (Ent) Formativa (Ent) Más que calificar (Ent)	El profesorado enfatiza la diferencia entre evaluar y calificar, donde el evaluar debe ser un proceso más complejo, que considere mayores ámbitos que el rendimiento. Se identifican indicadores en los instrumentos que valoran el proceso.
			Individual (Ent+Ins) Personalizada (Ent+Ins)	Se enfatiza el carácter individual de la evaluación, donde se debe adecuar a cada individuo y solo utilizar estándares generales. Los instrumentos son aplicables de forma individual.
			Individuallizada Intereses estudiante (Ent)	
			Adecuada (Ent)	
			Accesible (Ent)	

MARCO METODOLÓGICO

Prácticas evaluativas	Foco	Cualitativo	(Ent)	El profesorado manifiesta en la entrevista e instrumentos gran relevancia a las actitudes, comportamiento y esfuerzo. Por sobre el desempeño práctico, en menor medida lo conceptual.
		Esfuerzo	(Ent+Ins)	
		Comportamiento	(Ent+Ins)	
		Progreso	(Ent)	
		Conceptual	(Ent+Ins)	
		Procedimental	(Ent+Ins)	
		Actitudinal	(Ent+Ins)	
	Usos	Diagnóstico	(Ent)	Se mencionan diversos usos en el proceso, con énfasis en el seguimiento de actitud o comportamiento clase a clase, ya sea para calificar o no. Aunque la totalidad de instrumentos llevan calificación.
		Formativo	(Ent)	
		Calificar	(Ent+Ins)	
		Seguimiento de actitud	(Ent+Ins)	
		Observación clase a clase	(Ent+Ins)	
	Instrumentos	Registro clase a clase	(Ent+Ins)	Se mencionan y presentan diferentes instrumentos, los cuales deben estar validados por la Unidad Técnica Pedagógica (UTP) de cada establecimiento, bajo la norma del decreto 67.
		Escala de apreciación	(Ent+Ins)	
		Lista de cotejo	(Ent)	
		Rúbricas	(Ent+Ins)	
		Validados por UTP	(Ent)	
	Dimensiones	Actitudinal	(Ent+Ins)	Se visualiza un énfasis en la dimensión actitudinal, desde el comportamiento en clases, participación y esfuerzo por sobre las dimensiones procedimentales y en menor medida la conceptual. Similar proporción aparece en los indicadores de los instrumentos de evaluación.
		Comportamiento	(Ent+Ins)	
		Participación	(Ent+Ins)	
		Conceptuales	(Ent+Ins)	
		Esfuerzo	(Ent+Ins)	
		Superación	(Ent)	
		Procedimentales	(Ent+Ins)	

Nota. "Ent": Códigos provenientes de las entrevistas. "Inst": Códigos provenientes de los instrumentos de evaluación

3.3.4.2. Análisis estadísticos descriptivos

Con el propósito de complementar y contrastar los hallazgos cualitativos del Estudio 4, se incorporó un análisis estadístico descriptivo sobre las calificaciones oficiales de la asignatura de EFyS, reportadas por los establecimientos educacionales a nivel nacional. Las bases de datos utilizadas fueron proporcionadas por el Centro de Estudios del Ministerio de Educación de Chile, mediante una solicitud de acceso a la información pública, y contienen los promedios anuales de calificaciones desagregados por curso, nivel escolar, subsector y región.

El análisis se centró en el periodo comprendido entre los años 2018 y 2022, siguiendo criterios metodológicos y normativos previamente justificados. A partir de la totalidad de registros anuales reportados (ver estudio 4), se aplicaron los siguientes filtros: (a) por subsector, conservando únicamente la asignatura de EFyS; y (b) por nivel educativo, excluyendo los registros correspondientes a la modalidad de educación básica y media para adultos, a fin de mantener la comparabilidad con el sistema escolar regular.

Se utilizó estadística descriptiva para calcular medidas de tendencia central y dispersión, incluyendo la media, la desviación estándar y la moda de las calificaciones. Estos indicadores se calcularon tanto a nivel nacional como por región y por nivel educativo (básica y media). Para el procesamiento de los datos se empleó el software estadístico SPSS, versión 25.

Adicionalmente, se elaboró un histograma para representar gráficamente la distribución de los promedios nacionales correspondientes al año 2022, lo que permitió visualizar la concentración de las calificaciones reportadas y analizar patrones generales en la asignación de notas. Este enfoque permitió no solo describir el comportamiento evaluativo desde una perspectiva cuantitativa, sino también contrastarlo con los discursos del profesorado en ejercicio recogidos mediante entrevistas, fortaleciendo así la triangulación metodológica de la investigación. Todo lo mencionado anteriormente se puede observar en el apartado de resultados del estudio 4.

3.4. Consideraciones éticas de la investigación

Este estudio se desarrolló en estricto cumplimiento de las normativas éticas vigentes en investigación con personas. El proyecto fue aprobado por el Comité Ético Científico de la

Universidad Autónoma de Chile, según consta en el Acta N° CEC 38-21 de fecha 13 de diciembre de 2021 (ver Anexo 1). Esta aprobación certifica que el protocolo de investigación se ajusta a los estándares internacionales de ética en investigación, incluyendo los principios establecidos en la Declaración Universal sobre Bioética y Derechos Humanos (UNESCO, 2005), la Declaración de Helsinki de la Asociación Médica Mundial (2024) y la normativa ética propia del área de estudio.

De acuerdo con el dictamen del Comité de Ética, el estudio cumple con tres criterios fundamentales:

- No vulnera la dignidad de las personas participantes,
- Garantiza el respeto a los Derechos Humanos durante todo el proceso investigativo y
- Asegura la confidencialidad, el anonimato y la protección de los datos obtenidos.

Todos los procedimientos metodológicos fueron diseñados bajo principios éticos de respeto, justicia y autonomía. Se utilizaron consentimientos informados específicos para cada grupo participante: uno dirigido al estudiantado universitario (ver Anexo 2) y otro al profesorado en ejercicio (ver Anexo 3). Ambos documentos explicitan la voluntariedad de la participación, el uso confidencial de la información, el derecho a retirarse del estudio en cualquier momento y la finalidad exclusivamente científica del trabajo. Los y las participantes fueron debidamente informados sobre los objetivos del estudio, las actividades solicitadas, la duración estimada de su participación, los mecanismos de protección de datos personales y la inexistencia de beneficios o perjuicios directos asociados a su colaboración.

Todas las entrevistas y recolección de relatos se realizaron bajo condiciones de respeto y resguardo, siguiendo además las recomendaciones sanitarias vigentes al momento de la recolección de datos, y los registros fueron almacenados con medidas de seguridad y acceso restringido. Este compromiso ético no solo refleja el cumplimiento normativo, sino también una postura investigativa coherente con el paradigma cualitativo adoptado, que reconoce a las y los participantes como sujetos activos, reflexivos y merecedores de un trato ético en todas las etapas del proceso. Así, en la etapa 2 de la investigación, al solicitar los instrumentos

de evaluación posterior a la entrevista, no todo el profesorado lo facilitó para su análisis, lo cual fue respetado y explicitado en el análisis del estudio 3.

3.5. Criterios de validez considerados en la investigación

Es importante mencionar que esta investigación se enmarca por completo bajo un paradigma cualitativo, como ya se ha señalado anteriormente. En este sentido, la validez no puede entenderse desde los parámetros de la investigación cuantitativa, como la replicabilidad o la objetividad, sino que debe analizarse bajo criterios que respondan a la naturaleza interpretativa, contextual y situada del conocimiento que se produce (Lincoln y Guba, 1985).

No obstante, en la etapa final de la investigación, correspondiente al Estudio 4, se incorporó un análisis cuantitativo complementario a partir de datos secundarios de calificaciones oficiales en la asignatura de EFyS entre 2018 y 2022. Esta decisión respondió a una necesidad emergente derivada del trabajo cualitativo previo, en línea con un diseño secuencial exploratorio característico de enfoques mixtos (Creswell y Plano Clark, 2018). La inclusión de este componente cuantitativo no modifica la naturaleza cualitativa predominante del estudio, sino que la enriquece al aportar una triangulación metodológica que amplía y profundiza la comprensión del fenómeno evaluativo.

La validez cualitativa se sustenta en el compromiso ético y epistemológico por generar interpretaciones profundas, coherentes y plausibles del fenómeno investigado, considerando la complejidad de las experiencias humanas y la singularidad de los contextos en los que estas se producen (Sparkes y Smith, 2013; ; Smith, 2017). En este caso, así ocurre con el despliegue metodológico para llevar a cabo las dos etapas de esta tesis doctoral, dado que los objetivos buscan justamente interpretar experiencias, percepciones y prácticas del profesorado y estudiantado sobre la evaluación en EF. Desde este enfoque, la calidad de la investigación se juzga a partir de criterios como credibilidad, transferibilidad, dependibilidad, confirmabilidad y reflexividad (Smith y McGannon, 2018).

3.5.1. Credibilidad

La credibilidad alude a la fidelidad con que se interpretan las vivencias de los y las participantes desde su propio marco de referencia (Guba y Lincoln, 1989). En esta tesis, se implementaron diversas estrategias para fortalecerla en ambas etapas del estudio. En la primera etapa, los microrelatos permitieron a las y los estudiantes narrar libremente sus experiencias de evaluación en EF, mediante un formato escrito, anónimo y seguro (estudio 1). Posteriormente, los grupos focales (estudio 2) permitieron confrontar y ampliar dichas experiencias en una instancia colectiva de diálogo. En la etapa 2, con profesoras y profesores en ejercicio, se utilizó la entrevista semiestructurada, cuya guía fue validada por juicio de expertos (estudio 3), junto al análisis de los instrumentos de evaluación que el profesorado había utilizado recientemente en contexto real. Todo ello posibilitó una recogida de datos densa y situada. Además, se utilizaron citas textuales para representar fielmente la voz del estudiantado y del profesorado, una estrategia esencial para la credibilidad (Smith y McGannon, 2018).

3.5.2. Transferibilidad

La transferibilidad se refiere a la posibilidad de que los hallazgos de una investigación cualitativa puedan ser considerados significativos o útiles en otros contextos educativos similares, no mediante generalizaciones estadísticas, sino a través del reconocimiento de experiencias compartidas (Erickson, 1986; Smith, 2017). Tal como plantea Smith (2017), esta forma de generalización se basa en que las y los lectores puedan verse reflejados en las situaciones descritas, identificando aspectos que resuenen con sus propias realidades. Para favorecer este tipo de comprensión, resulta fundamental ofrecer una descripción rica y detallada del contexto, de las trayectorias de las personas participantes y de las condiciones educativas en las que se desarrollaron los estudios.

En esta investigación se buscó garantizar dicha profundidad descriptiva, incluyendo información sobre el contexto local, las carreras de formación del profesorado, el nivel académico, y las experiencias escolares y universitarias del estudiantado (estudios 1 y 2), así como del profesorado en ejercicio (estudio 3). Además, se consideró la diversidad de trayectorias y experiencias, atendiendo a variables como el género, el tipo de establecimiento

y las prácticas evaluativas vividas, lo que amplía las posibilidades de que los hallazgos sean valorados como pertinentes en otros contextos. Como señala Smith (2017), esta diversidad favorece una forma de generalización basada en las múltiples dimensiones de la experiencia social, lo que permite a otras y otros investigadores y docentes identificar similitudes y reflexionar sobre la aplicabilidad de los resultados en sus propias realidades educativas.

3.5.3. Dependibilidad

La dependibilidad refiere a la estabilidad lógica del diseño metodológico, asegurando que las decisiones tomadas sean comprensibles y trazables (Guba y Lincoln, 1989). En esta tesis, cada estudio fue descrito con precisión, desde los criterios de inclusión hasta los procedimientos de análisis. El guion de grupo focal, por ejemplo, fue diseñado a partir de los relatos escritos (estudio 2), evidenciando un proceso coherente y progresivo. Se documentaron todos los pasos, desde la recolección de datos hasta la codificación inductiva, lo que permite reconstruir el itinerario metodológico y verificar su consistencia (Sparkes, 2001). Todo lo mencionado anteriormente partió con un análisis de la literatura y estado del arte existente sobre evaluación en EF.

3.5.4. Confirmabilidad

Este criterio busca asegurar que los resultados provienen de los datos y no de las suposiciones del investigador principal (Jootun et al., 2009; Olmos-Vega et al., 2022). En esta investigación se emplearon múltiples estrategias, desde contar con un equipo de investigadores expertos, hasta la triangulación metodológica (relatos, grupos focales, entrevistas, documentos), uso de citas textuales, validación externa de instrumentos y matrices de análisis sistematizadas. En cada uno de los estudios presentados, se describe el procedimiento de recolección de información y se detalla el proceso de análisis de datos, donde las categorías eran emergentes de las unidades hermenéuticas obtenidas de estudiantado (etapa 1) o del profesorado (etapa 2).

3.5.5. Reflexividad

La reflexividad implica reconocer cómo la posición del investigador/a influye en todo el proceso investigativo (Finlay, 2002). En esta tesis, se explicita que el equipo investigador está compuesto por docentes y formadores en EF, lo que supone una implicación directa con

el objeto de estudio. Se promovió una postura reflexiva durante todo el proceso, desde el diseño hasta la interpretación, incorporando espacios de codificación colaborativa y discusión con otros investigadores. Este ejercicio ético y epistemológico es clave para mitigar sesgos y potenciar la riqueza interpretativa (Gerstl-Pepin y Patrizio, 2009; Townsend y Cushion, 2021). Un ejemplo claro que permitió desestimar un sesgo fue el considerar de forma equitativa la distribución por género y por especialidad de pedagogía en la primera etapa.

En resumen, estos criterios no solo fortalecen la validez de la tesis doctoral, sino que dan cuenta de un proceso investigativo éticamente comprometido, rigurosamente documentado y epistemológicamente coherente con los principios de la investigación cualitativa. La articulación entre técnicas, fuentes y voces permite ofrecer una mirada compleja, situada y significativa sobre las prácticas y experiencias evaluativas en EF, tanto desde la perspectiva del estudiantado como del profesorado en ejercicio.



VNIVERSITAT E VALÈNCIA

CAPÍTULO 4

ESTUDIO 1

Estudio 1. Dime cuánto corres y te diré qué nota mereces. Evaluación en la Educación Física chilena.

Estado: Artículo aceptado en la revista Cultura, Ciencia y deporte (CCD).

Resumen

La importancia de la evaluación como elemento definitorio de todo proceso educativo genera la necesidad de indagar sobre experiencias del futuro profesorado chileno en su etapa de formación escolar, identificando significados, sentimientos y consecuencias que generaron los test de condición física usados en la asignatura de Educación Física. A partir del estudio de 60 relatos narrativos de estudiantes de diferentes especialidades de Pedagogía, se generó un análisis categorial inductivo. Los resultados manifiestan alta prevalencia de los test de condición física para evaluar y calificar la asignatura, provocando estrés y sensación de ansiedad. El estudiantado con bajos niveles de condición física manifestaba sentimientos de frustración, desmotivación e inseguridad, hasta el extremo de desertar de la clase mediante justificantes o excusas diversas. Por otro lado, quienes declaran buenos niveles de condición física y cuyas aptitudes encajaban con los estándares exigidos, los test generaban motivación y satisfacción. Como conclusión, los test de condición física utilizados en Educación Física para calificar conllevan una gran carga emocional para todo el estudiantado, especialmente para quienes no presentan altos niveles de condición física. En estos casos prevalecen sentimientos de frustración y malestar hacia la asignatura por las experiencias negativas que esta genera en el alumnado.

4.1. Metodología

4.1.1. La investigación narrativa

Esta investigación cualitativa emplea relatos biográficos enmarcados en la investigación narrativa (Bolívar, 2002) para recoger las vivencias del alumnado relacionadas con la evaluación en la asignatura de Educación Física y Salud. La metodología se fundamenta en la hermenéutica y la fenomenología, permitiendo interpretar las experiencias y sentimientos expresados por los participantes (León, 2015).

La investigación narrativa, que ha ganado fuerza en el área de la EF (Devís-Devís, 2017; Dowling et al., 2013; Groves y Laws, 2003;), ofrece la posibilidad de adaptarse al contexto investigado y recoger los relatos de manera abierta y reflexiva (Kincheloe, 2001; Pérez-Samaniego et al., 2011). En este estudio, los relatos biográficos permiten comprender cómo el alumnado interpreta sus experiencias escolares relacionadas con la evaluación, proporcionando un marco adecuado para el análisis narrativo.

4.1.2. Participantes

En la presente investigación participaron 60 estudiantes de la Facultad de Educación de una universidad de la región del Maule en Chile, pertenecientes a 6 grados diferentes: 10 de Pedagogía en Lengua Castellana y Comunicación, 10 de Pedagogía en inglés (PI), 10 de Pedagogía en Educación Física (PEF), 10 de Pedagogía en Educación Parvularia (PEP), 10 de Pedagogía General Básica (PGB) y 10 de Pedagogía en Historia y Geografía (PHG). En el instrumento se consultó por el género, dando las opciones de femenino, masculino y no binario, resultando la distribución que se presenta en la Tabla 17.

Tabla 17. Distribución según género de las personas participantes.

Disciplina	Femenino	Masculino	Cantidad
Pedagogía en Lengua Castellana y Comunicación	7	3	10
Pedagogía en Educación Física	3	7	10
Pedagogía en Inglés	5	5	10
Pedagogía en Educación Básica	7	3	10
Pedagogía en Historia y Geografía	3	7	10
Pedagogía en Educación Parvularia	10	0	10
Total	35 (58%)	25 (42%)	60 (100%)

Elaboración propia

En base a su experiencia y formación en Pedagogía en el año 2022, se invitó a participar al alumnado del nivel 5 de cada grado, lo que corresponde con el primer semestre del tercer año de su formación en el cual, según el plan curricular, todos y todas deberían tener formación en la asignatura de didáctica y evaluación. Por tanto, la edad de los y las participantes oscilaba entre los 20 y los 30 años con un promedio de 22,41 y una desviación estándar de 3,15.

Dentro de los criterios de inclusión, se tuvo en cuenta que el estudiantado hubiera vivenciado la asignatura de EFyS en su edad escolar de forma regular, es decir, sin haberse eximido de ello durante un periodo mayor a un año.

4.1.3. Procedimiento

La invitación a participar se hizo en cada uno de los grados para explicar el contexto y objetivo de la investigación. El estudiantado accedía a participar firmando un consentimiento informado aprobado por el comité de ética de la Universidad participante (Acta N° Acta N°CEC 38-21). Una vez firmados los consentimientos, se procedió a reunir al grupo en un aula específica, en la cual accedieron a un formulario virtual para responder a preguntas breves asociadas a una caracterización general y, finalmente, llegar al encabezado del relato narrativo que se les estaba solicitando. Todas las personas accedieron al relato al mismo tiempo, el cual fue leído en voz alta por el investigador principal. El encabezado utilizado fue el siguiente:

“En esta etapa es necesario que te remontes a tu edad escolar, recordando vivencias desde primero básico a cuarto medio, con respecto a la EVALUACIÓN en la asignatura de EF o EFyS. La idea es que nos relates, en base a tus experiencias reales, los procesos de Evaluación que vivenciaste en la asignatura. El estilo de relato depende de tu comodidad, puedes escribir un relato general o momentos específicos que nos ayudarán a construir la realidad que experimentaste. Es muy relevante, además, describir cuales fueron tus percepciones sobre esas instancias de Evaluación en la asignatura, qué sentimientos fueron los que generaron esas experiencias, ya sea desde el agente evaluador, instrumento, clima, sentido, etc.”

El encabezado usado en esta investigación otorgaba libertad a los y las participantes para relatar sus experiencias y vincularlas emocionalmente según los puntos que ellos o ellas consideraran relevantes.

Como criterios éticos y de credibilidad, se utilizaron diversas estrategias que se exponen a continuación: a) el encabezado y todo el procedimiento fue elaborado conjuntamente por el investigador principal junto a un equipo internacional con amplia experiencia en investigación cualitativa, evitando una propuesta limitada a la mirada de un solo investigador;

b) la participación del alumnado fue totalmente voluntaria y anónima; c) el investigador que aplicó el instrumento no era docente directo de ningún participante del estudio, como hicieron antes Cornejo y Salas (2011); d) se llevaron a cabo dos reuniones posteriores con los participantes, donde se presentaron las interpretaciones y principales resultados, siendo avalados por los propios protagonistas (del Villar-Álvarez, 1994).

4.1.4. Análisis de datos

Si bien existen diferentes estrategias para generar un análisis narrativo desde el cual comprender lo que la voz de los y las protagonistas quiere expresar, se optó por el análisis categorial inductivo (Hodge y Sharp, 2016; Pérez-Samaniego et al., 2011). Esta estrategia deja ‘hablar’ a los datos obtenidos ya que cada participante del estudio pudo escribir libremente respecto de sus experiencias en evaluación. El proceso de análisis comenzó con la identificación de las unidades de significado explícitas, asignando un código a cada una de ellas. Una vez finalizada esta primera etapa de consulta constante por parte del equipo de investigación se procedió, como explican Trainor y Bundon, (2021), a organizar y categorizar la larga lista de códigos que había emergido del primer análisis, proceso que redujo bastante las categorías. En la tabla 18 se presenta la matriz de categorías que emergieron tras el análisis inductivo en la que el número entre paréntesis se corresponde con la frecuencia de aparición del código.

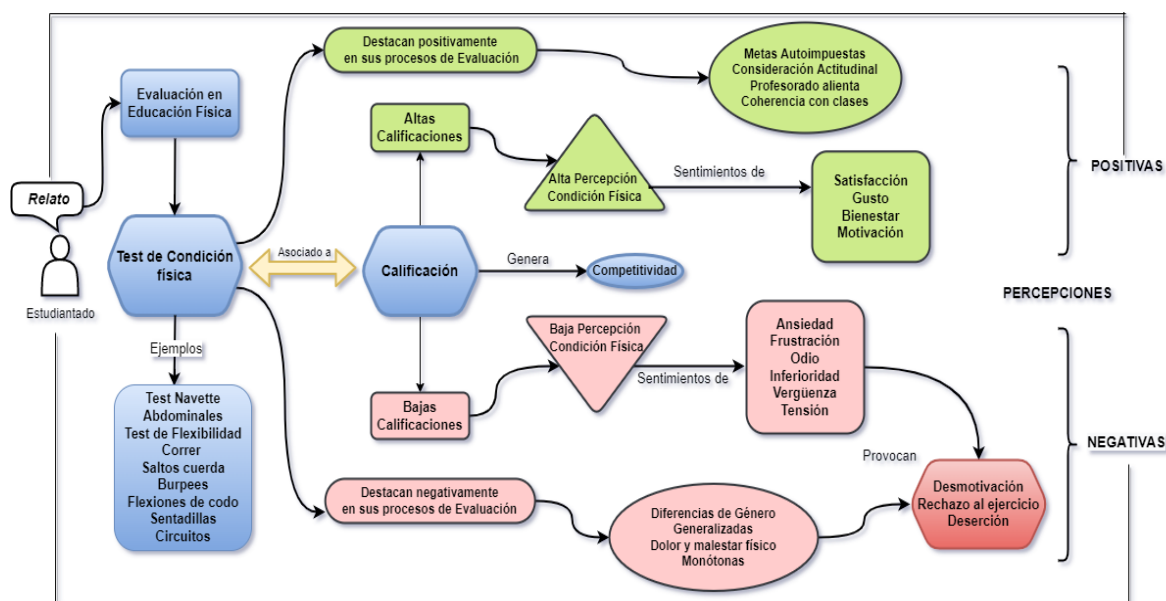
Tabla 18. Matriz de categorías emergentes de los relatos narrativos.

Categoría	Ejemplos mencionados	Subcategoría	Códigos percepciones
Experiencias con Test de Condición Física (60)	Test Course Navette (27) Abdominales (14) Correr (14) Flexibilidad (5) Salto de Cuerda (3) Test de Cooper (3) Salto Largo (2) Burpees (2) Circuitos (2)	Manifestaron baja percepción de la condición física	Frustración (10)
			Desmotivación (8)
			Enojo (6)
			Vergüenza (6)
			Malestar físico (6)
			Inseguridad (5)
			Desgaste físico (3)
			Inferioridad (3)
			Satisfacción (10)
		Manifestaron alta percepción de la condición física	Superación (5)
			Motivación (2)
			Gusto (2)

Elaboración propia

Dado el carácter narrativo de los relatos y la diversidad de experiencias descritas, se diseñó un mapa conceptual (Figura 5) para organizar las temáticas emergentes y trasladar los datos desde lo inductivo a lo general (Nowell et al., 2017). Este proceso permitió estructurar las categorías finales en función de diversos itinerarios en cuatro grandes dimensiones: (1) "Dime cuánto corres y te diré qué nota mereces", en referencia a las evaluaciones y calificaciones basadas en el rendimiento en test de condición física; (2) "Alto desempeño, altas notas, alta motivación", asociada a experiencias positivas de estudiantes con buena percepción de su condición física; (3) "Bajo desempeño, bajas calificaciones, baja motivación", relacionada con percepciones negativas vinculadas a dificultades experimentadas en estas pruebas; y (4) "El abandono", como consecuencia del impacto emocional negativo derivado de estas evaluaciones. Estas categorías finales surgieron tras la agrupación de códigos y la identificación de patrones comunes en los relatos. Para reforzar la relación entre los datos y su interpretación, se incorporaron citas textuales siguiendo una nomenclatura estructurada, lo que permitió evidenciar cómo los discursos del estudiantado sustentan las categorías identificadas.

Figura 5. Mapa conceptual sobre las experiencias en evaluación con test de condición física.



Elaboración propia

Posteriormente, como alternativa al formato académico tradicional, se utilizaron técnicas literarias para introducir citas textuales del estudiantado, utilizando un número de

identificación, luego las iniciales del grado y la última letra referente al género (ejemplo 2 PI F; que corresponde al segundo estudiante de Pedagogía en inglés de género femenino). Además, las citas textuales ayudan a exponer interpretaciones y a refrendar los resultados. Para ello, se siguieron las recomendaciones de cómo representar datos empíricos de forma creativa (Smith et al., 2015).

4.2. Resultados

A modo de contextualización, en términos generales el estudiantado manifestó en sus relatos acciones evaluativas asociadas siempre a la calificación, por lo que en los resultados se mencionan como símiles, dejando ver un enfoque tradicional para dar cumplimiento a las normativas de calificación en cada contexto. A su vez, se manifestaron vivencias en la asignatura correspondientes en mayor medida a estilos de enseñanza tradicionales y reproductivos, similares en ocasiones a instrucciones militares. Dicho estilo presentó matices diferentes en las metodologías utilizadas en deportes o en danzas folklóricas. Para esta investigación, se analizaron los relatos en base al uso de test de condición física en situaciones de evaluación vinculadas a la calificación.

Dime cuánto corres y te diré qué nota mereces

Aunque correr sea la principal metáfora para explicar el sistema de evaluación hegemónico, no es la única actividad física que parece contar a la hora de ‘sacar’ una calificación u otra en EFYS. En realidad, también podría usarse la distancia de salto, el número de sentadillas o las abdominales que eran capaces de realizar en un determinado tiempo:

Todas las evaluaciones que he realizado en mi vida sobre educación física fueron sobre resistencia física (4 HG F).

La profesora nos evaluaba el hacer flexiones de brazo, abdominales y saltar hasta una cierta distancia, donde éstas tres notas se promediaban y daban una sola calificación y era de acuerdo con lo que uno hacía (9 EP F).

Las evaluaciones iban orientadas a un alumno ideal. Recuerdo que el profesor tenía una tabla donde salía cuánto debía saltar cada estudiante, cuánto debía durar en el Test Course Navette o cuántos abdominales/flexiones debería hacer por minutos. Cuando vi eso, sentí que estaba muy lejos de ese ideal (8LC M).

La totalidad de los y las 60 estudiantes manifestaron haber vivenciado experiencias evaluativas con test de condición física en su época escolar asociadas a una calificación. Dentro de las pruebas más mencionadas están la Course Navette, hacer abdominales, test de flexibilidad, correr una distancia en un tiempo determinado, saltos de cuerda, *burpees*, flexiones de codo, sentadillas o circuitos. Así, la calificación se generaba únicamente por el rendimiento en la prueba física respectiva. Cuando el desempeño era menor al esperado, la calificación bajaba según tablas que el profesorado, en ocasiones, entregaba, aunque no siempre:

En las herramientas de evaluación, para la primera unidad sólo se guiaba por los estándares establecidos, acorde a las marcas correspondientes según el género y edad de cada estudiante (4 EF F).

Otra de las formas en que el docente nos evaluaba era dando vueltas por toda una manzana por afuera del liceo. Cada vez que completábamos una vuelta la profesora nos marcaba una línea en la mano con un plumón, dependiendo de la cantidad de líneas en nuestras manos, era nuestra nota (7 LC F).

En la enseñanza básica se practicaba “el famoso” test Course Navette, que al llegar a una cierta cantidad de minutos se cumplía con la nota 7,0, que es la máxima. Yo siempre obtenía el 7,0, sin embargo, daba hasta el último respiro por conseguir el objetivo o sería mal calificada. En ese aspecto me veía muy presionada ya que, al no rendir físicamente, podría perjudicar mi promedio en general (3 LC F).

La competitividad y la ansiedad asociada a estas pruebas eran factores comunes entre el alumnado, cuyo fin no era otro que calificar y clasificar el rendimiento individual de cada estudiante. Así, incluso cuando se tenía una buena estima hacia la asignatura, la evaluación generaba sentimientos de nerviosismo:

Al finalizar cada período venía una evaluación y generalmente estas eran de carácter práctico, en donde uno tenía que correr, saltar, estirarse, trotar, Course Navette, etc. Generalmente, esto me ponía algo ansioso porque se generaba un ambiente de competencia entre compañeros (4HG M).

En la enseñanza media me empezó a gustar más lo que era la actividad física, pero cuando llegaba el momento de evaluar, era un momento que me ponía nerviosa por querer llegar a la meta y poder obtener la más alta calificación (9 EP F).

Alto desempeño, altas notas, alta motivación

Del análisis de los relatos se desprende que solo una minoría del alumnado lograba cumplir con las expectativas de los test, ya sea por tener una buena condición física o bien por ser de su agrado una metodología militarizada:

Ya que era considerada una de las "buenas" para el deporte, las evaluaciones no tuvieron una mayor dificultad que la mínima, ya que podía hacerlas sin problema (10 PB F).

Otra de las evaluaciones que nos realizaba la profesora, eran los abdominales, me gustaba mucho poder llegar a la meta y cumplir con lo que la profesora nos exigía (4 PB F)

Eran muy intensas, similares a como cuando adoctrinan a algún militar, pero para mí no era problema ya que me gustaba cuando me exigían y yo me exigía más (5 HG M).

Los sentimientos de satisfacción, gusto y bienestar por los test de condición física se relacionan con el cumplimiento de las expectativas depositadas en cada cual, siempre en los casos de rendimiento óptimo y, por tanto, generando calificaciones máximas. En estos casos, se manifiesta una alta motivación hacia la asignatura y el sistema de evaluación planteado:

En la básica me encantaba la evaluación del Test Course Navette, ya que siempre me iba muy bien y tenía muy buen rendimiento, amaba que llegara el día jueves y tener que correr y hacer diferentes actividades (4PB F).

En menor medida, una valoración positiva hacia los test de condición física se generaba cuando el profesorado consideraba la dimensión actitudinal dentro de su evaluación, o cuando se manifestaba una coherencia de las clases con la forma de evaluar:

Solo evaluaban nuestro rendimiento físico, aunque sí debo rescatar que involucraban indicadores de evaluación de dimensiones actitudinales, que por cierto es característico de la Educación Física (9 EF M).

Así, por ejemplo, el simple hecho de animar al estudiantado durante el esfuerzo ya suponía un cambio digno de ser recordado:

Nos hizo realizar un test el cual suponía dar 16 vueltas a las 3 canchas que había, en 12 minutos. Yo, cuando iba en la décima vuelta, no podía más, estaba casi con la lengua fuera y él me miró y cada vez que yo pasaba cerca de él, me animaba y no tan solo a mí, sino que a todos los que estábamos rindiendo ese test, era alguien que era capaz de cambiar la vida de alguien (6 EF M).

Bajo desempeño, bajas calificaciones, baja motivación

Cuantitativamente hablando, más de la mitad de los relatos analizados recogían sentimientos de ansiedad, frustración, odio, vergüenza e inferioridad al verse enfrentados a este tipo de escenarios evaluativos. Estos sentimientos negativos se generaban al no cumplir con las exigencias físicas de las diferentes pruebas por bajos niveles de condición física:

Mi experiencia no fue tan grata pues el profesor me exigía mucho rendimiento físico y yo no tenía tanta resistencia física, entonces me frustraba no poder alcanzar la meta de los ejercicios y me enojaba (10 PB F).

Siempre fue una gran batalla responder de forma correcta a todas las actividades de Educación Física, porque siempre fui la más pequeña y a todos nos exigían el mismo rendimiento. Siempre fue frustrante ya que, para mí, las calificaciones eran importantes y nunca lograba la exigencia que se me pedía (7 EP F).

Dicha desmotivación y rechazo hacia la asignatura se vio potenciada por el malestar físico y sufrimiento causado por el esfuerzo de lograr, no ya una calificación máxima, sino llegar al mínimo exigido. En estos casos, los recuerdos de odio hacia este tipo de test son los que prevalecen en los recuerdos de la asignatura:

Nos evaluaba a través del Test Course Navette, recuerdo que fue una de las peores experiencias que he vivido porque estaba muy cansada y sentía que me faltaba el aire, y sentía mi corazón latir muy fuerte, me mareé y tenía ganas de vomitar, a parte que era un test que dependiendo de la cantidad de tiempo te evaluaban, recuerdo también que tenía compañeras con sobrepeso y ellas no podían dar mucho y obtenían una mala calificación por ello (7 LC F).

Recuerdo lo que sufría con la odiada prueba Course Navette, además le tenía un poco de temor a la profesora. Corría hasta cuando no podía respirar, me sentía tan débil, llegando a dolores de cabeza para sacarme una nota mejor y lo peor de todo, es que a mí no me iba tan mal en comparación con mis compañeros de curso (3 EP F).

El abandono

Las bajas calificaciones obtenidas tenían como resultado directo sentimientos de frustración, lo que conllevó en varios casos a desertar de la asignatura. Esto se conseguía con justificantes inventados para tratar de evitar repetir las pruebas de evaluación:

Es que veía como compañeros se entristecían al no poder alcanzar las expectativas del profesor y de los propios compañeros, por lo que se ponían a llorar o sencillamente no rendían la prueba y se ocultaban bajo cualquier justificativo (3HG M).

Pero bueno, luego de conocer este test, empezó el famoso cuento de que andaba en mi periodo (menstrual), cada vez que nos tocaba rendir el Test Course Navette para tratar de no hacerlo, ya que no quería vivir de nuevo este tipo de vergüenza (8PB F).

Este absentismo en la asignatura provocaba un sentimiento de alivio al no tener que pasar por las evaluaciones, las cuales causaban gran malestar y presión psicológica, disminución de su autoestima y sentimiento manifiesto de renunciar de forma permanente a las clases de EFYS:

En educación media opté por no hacer nada... evitaba a toda costa sentirme sudada o "torpe". (4 LC F).

Con esto es que tomé la decisión de ir con falda al colegio para no realizar actividad física, por las características del colegio, Educación Física solo se realizaba en 1° y 2° medio. En tercero medio sentí un alivio con respecto a que ya no me obligarían a realizar 50 sentadillas en dos minutos o que solo se me evaluara por mi resistencia física (4 HG F).

4.3. Discusión

Este estudio aborda las consecuencias e interpretaciones que tiene para el alumnado el uso de los test de condición física en la asignatura de EF en su etapa escolar, recordando sus vivencias desde una perspectiva pedagógica en base a su formación como futuro profesorado de las diferentes disciplinas consideradas.

Si bien en el programa chileno de la asignatura de EF existe un indicador evaluativo para el que se sugiere la realización de un test de condición física, el mismo programa recomienda no calificar en función del rendimiento en dichos test y sugiere utilizarlo, solo, con fines formativos (Ministerio de Educación, 2016). Sin embargo, la realidad sigue mostrando como

se perpetúan dichos test de condición física para calificar, valorando la competencia motriz y los aspectos procedimentales de la misma, dando paso a la discriminación en el aula (Moreno-Doña et al., 2016; Junior et al., 2022; Rodríguez-Rodríguez et al., 2017). De hecho, este tipo de pruebas estandarizadas establecen una lógica evaluativa escasamente formativa, llegando incluso a la desvalorización profesional de los y las docentes (Moreno-Doña et al., 2014; Oliver-Álvarez y Martos-García, 2023).

El fin de la evaluación, en general, y de los propios test de condición física, en particular, debería estar relacionada con la mejora de los procesos de enseñanza y de aprendizaje, enfatizando los aspectos formativos a lo largo del proceso educativo, tal como lo señala el decreto 67 (Ministerio de Educación, 2018) que, además de regir la calificación y promoción, establece directrices generales sobre el sentido formativo que se debería dar. No obstante, otorga la responsabilidad a cada centro educativo para que propongan sus propios reglamentos de evaluación. Con todos los lineamientos curriculares y legislativos mencionados anteriormente que dotan de cierta libertad al profesorado de proponer diversas instancias de evaluación, la medición y el rendimiento del alumnado todavía ocupan un lugar importante entre las finalidades atribuidas a la misma, como hemos visto, lo que coincide con los argumentos de Olague y Cañadas (2021).

Es relevante mencionar que, en las narraciones de las vivencias sobre los test de condición física de su edad escolar, el futuro profesorado de EF destacó positivamente la inclusión de elementos actitudinales entre los criterios de evaluación, lo que refleja su conocimiento específico de la disciplina (Rodríguez et al., 2018; Silva et al., 2024). Y es fundamental que se promueva esta práctica, no solo en la evaluación, sino desde la propia metodología abordada en las clases y que contempla el currículum (Mujica Johnson, 2019).

Cabe señalar que los relatos que manifiestan experiencias positivas frente a los test de condición física en la asignatura de EF fueron notoriamente menos frecuentes. Los sentimientos de satisfacción por cumplir las expectativas que mencionaron quienes relataron tener una buena condición física, se pueden explicar por el gusto de obtener buenas calificaciones o bien por la motivación generada al obtener refuerzos positivos y cierto reconocimiento (Álvarez-Rivera et al., 2015; Nabaskues et al., 2019). Cabe señalar que, en algunos de estos relatos, al mismo tiempo que relatan sus experiencias positivas con los test

de condición física, recuerdan con preocupación a sus compañeros con sobrepeso u obesidad, a quienes les resultaba más difícil cumplir con las metas propuestas por el profesorado. Este ejercicio de empatía se da en otros estudios anteriores a este (Bores-García et al., 2021; Mújica-Johnson et al., 2017; Oliver et al., 2021).

Dicho esto, y de acuerdo al carácter narrativo de esta investigación, merece la pena poner de manifiesto como las interpretaciones expresadas por el estudiantado chileno sobre sus experiencias en las actividades de evaluación en EFyS concuerdan bastante con los resultados de otras investigaciones (Danthony et al., 2020; Nabaskues et al., 2019; Phillips et al., 2017). Por ejemplo, destacan la presión y sensación de ansiedad que se dan en situaciones concretas de calificación orientadas al desempeño físico (Calatayud, 2021; Sánchez-Hernández et al., 2022). Este hecho se puede explicar por la importancia de la ponderación de las calificaciones que dan acceso a la educación superior ya que, en el contexto chileno, el promedio de calificaciones (en el cual se incluye la asignatura de EFyS), puede influir en la elección del grado que se desea estudiar, o los beneficios que se pueden obtener, inclusive la gratuidad de estos estudios (Ministerio de Educación, 2023). Otros trabajos han corroborado la influencia que ejercen las calificaciones en la forma que tiene el alumnado de experimentar su proceso formativo (Martos-García y García-Puchades, 2023).

Dada la importancia que tienen las calificaciones en el sistema escolar, además de la presión mencionada anteriormente, surge el sentimiento de competitividad que, para algunos, era una motivación y generaba un deseo por superarse, pero para otros suponía momentos de tensión dentro del ambiente del curso. La literatura confirma lo expuesto por el estudiantado en donde la competencia percibida se puede relacionar negativamente con la ansiedad en caso de personas con una condición física menor (Baghurst y Mwavita, 2014; Danthony et al., 2020; Wrench y Garrett, 2008) y positivamente con el compromiso con el aprendizaje o la actividad realizada (Coterón et al., 2013). La literatura también relaciona una autopercepción negativa con sentimientos de ansiedad física y social (Camacho-Miñano et al., 2006; Escalante-Izeta et al., 2016).

Los sentimientos de frustración, enojo y vergüenza emanan de las situaciones relatadas en las que se calificaba al grupo por medio de un test de condición física, donde el único criterio de evaluación era cumplir con los estándares establecidos por el profesorado de la

asignatura. Así, las interpretaciones dependen de las características personales de quienes lo vivenciaron, potenciando las emociones negativas entre las personas físicamente inactivas (Beltrán-Carrillo y Devís, 2019). Cuando las metodologías y modelos pedagógicos de la clase destacan aspectos actitudinales, se atribuyen emociones positivas (Bores-García et al., 2021; Mújica-Johnson et al., 2017), pero cuando se relacionan con el proceso de evaluación sumativa por medio de test de condición física, sumados a una baja percepción de la propia condición física, las emociones se tornan negativas no solo para el estudiantado, sino que también viene a tensionar un proceso que para el profesorado genera ansiedad y dificultad (Sonlleve et al., 2018; Véliz et al., 2020), como lo es el diseñar y proponer vivencias evaluativas que promuevan la trascendencia de un aprendizaje significativo.

Cuando se evocan recuerdos de malestar físico, dolor y molestias asociadas a la evaluación de la EF por medio de test de condición física, se relacionan principalmente con el test Course Navette, que es una prueba utilizada para conocer el consumo de Oxígeno Máximo de un individuo por medio de ecuaciones según la edad o el género, entre otras variables. Es utilizada principalmente por su bajo coste en implementación y una alta correlación comparada con un método directo (López-Gil et al., 2020). Al ser considerado un test que supone un máximo esfuerzo, y considerando la escasa frecuencia de la propia clase (una clase por semana en general) o el estilo de vida del escolar chileno, se podría considerar una actividad de riesgo sin una correcta supervisión (Curilem et al., 2012). Por otra parte, el llegar a la fatiga y dolor muscular se asocia a emociones negativas (Dhantony et al., 2020; Hugo et al., 2021), lo que explicaría gran parte de los sentimientos generados por estas pruebas y la necesidad por conseguir una mejor calificación.

El deseo de abandono de la clase, o buscar excusas para no realizar las evaluaciones, se explica por motivos similares a la deserción deportiva, donde las experiencias negativas y la desmotivación promueven el deseo de alejarse de todas estas experiencias de sufrimiento (Gil et al., 2012). En los relatos se manifestó, de forma textual, que lo vivido en la asignatura puede considerarse una de las peores experiencias a recordar de la edad escolar, como han argumentado antes Oliver et al. (2021). Dicho abandono y el consecuente rechazo a la práctica de actividad física intensa tienen pronósticos lamentables para la realidad chilena en relación con la condición física y la salud (Aguilar-Farias et al., 2016).

Para finalizar, las recomendaciones entregadas por los sujetos de estudio en sus relatos hacen énfasis en no asociar estas prácticas poco pedagógicas a una calificación y sin un sentido formativo, tal como lo viene reforzando la literatura desde hace décadas (Mora, 2004). Los modelos de evaluación tradicionales con un enfoque técnico van quedando desplazados por propuestas como la evaluación auténtica (Ahumada, 2005) o la formativa y compartida (López-Pastor y Sicilia-Camacho, 2017), en donde las experiencias y resultados señalan que el alumnado percibe un mejor clima en general dentro de las sesiones de clase, lo que contribuye sensiblemente a su aprendizaje en la materia y a su rendimiento académico (López-Pastor et al., 2012; Silva et al., 2024). Dichas interpretaciones de la evaluación y del currículum se deberían abordar con mayor énfasis en la formación inicial docente que prepara al futuro profesorado de EF (Asenjo-Paredes et al., 2024).

4.4. Conclusiones estudio 1

Los test de condición física siguen estando fuertemente presentes en la asignatura de EFYS en el contexto chileno, evocando contundentes emociones que alejan al estudiantado de la práctica de la clase de EF.

Al asociar dichas pruebas a una calificación por desempeño físico se generan sentimientos de ansiedad y competitividad. Cuando el estudiantado presenta una buena condición física, la interpretación de los test de condición física genera sentimientos de satisfacción, bienestar y motivación por el hecho de alcanzar los estándares solicitados por el profesorado, aunque se manifiestan empáticos por sus compañeros y compañeras que sufrían en dichos escenarios de evaluación.

Por otra parte, cuando el estudiantado manifiesta bajos niveles de condición física, alejándose de las expectativas del profesorado, los sentimientos que afloran son de frustración, ansiedad, odio, inseguridad, vergüenza y malestar, lo que generaría el rechazo y deseos de abandonar la asignatura, llegando incluso a inventar excusas para faltar a clases. También evocan recuerdos de malestar y monotonía, generando sentimientos de rencor hacia la asignatura y el ejercicio físico.

Todo lo anterior se contrapone a la necesidad de promover mediante la clase de EFYS un estilo de vida saludable asociado a la actividad física y al cuidado de la salud, que es uno de los propósitos del curriculum nacional de la asignatura.

Finalmente, se sustenta lo recomendado por las bases curriculares de la asignatura; los Test de condición física no deberían utilizarse para calificar al estudiantado por su rendimiento o estado físico, sino más bien con fines formativos.

Como principal limitación del estudio cabe resaltar la gran variedad de temáticas abordadas en los relatos, ya que al dar la libertad de expresar sus vivencias de forma abierta se generó una enorme cantidad de información asociada a las metodologías, características del profesorado de EF y recomendaciones en base a su propia formación como futuro docente.

Para complementar esta investigación sería interesante conocer la versión del profesorado que continúa perpetuando la utilización de test de condición física para generar calificaciones en diferentes establecimientos, profundizar en sus razones y en los sentidos desde los que se usan los modelos tradicionales de evaluación.

Finalmente, dar voz al estudiantado, para evitar replicar este tipo de prácticas, entrega un sustento epistemológico que podría ser utilizado en la formación inicial docente del profesorado de EF.



VNIVERSITAT DE VALÈNCIA

CAPÍTULO 5

ESTUDIO 2

Estudio 2. Evaluación en Educación Física escolar y su impacto en la percepción de la asignatura: vivencias narrativas de estudiantes de pedagogía.

Estado: Artículo en revisión.

Resumen

Este estudio cualitativo explora cómo las experiencias evaluativas en Educación Física influyen en las actitudes y percepciones del estudiantado sobre la propia asignatura. Utilizando relatos narrativos y grupos focales, se analizaron las vivencias de 60 estudiantes para comprender cómo las prácticas evaluativas afectan la adherencia y la percepción hacia la asignatura. Los resultados revelan que la evaluación mediante test físicos y actividades deportivas estandarizadas, como el Course Navette y la gimnasia, genera tanto satisfacción en aquellos estudiantes que logran cumplir con los estándares, como frustración y sentimientos de injusticia en quienes no alcanzan dichas expectativas. Este efecto es particularmente evidente en situaciones de discriminación donde prevalece el sexismo y la "gordofobia". Además, se destaca la ausencia de explicaciones pedagógicas claras sobre el propósito de las evaluaciones y la falta de transparencia en los indicadores utilizados, lo cual contribuye a una percepción de discrecionalidad en las calificaciones. Las evaluaciones de danza fueron percibidas de manera diversa: algunas personas destacaron su carácter colaborativo, mientras que otras las percibieron como obligatorias y monótonas. En general, se observó una actitud de desmotivación hacia la asignatura, estrechamente vinculada con un enfoque evaluativo que no fomenta el aprendizaje significativo. En conclusión, este estudio subraya la necesidad de replantear las prácticas evaluativas, promoviendo enfoques más inclusivos y justos que consideren la diversidad del alumnado.

5.1. Metodología

Esta investigación se ubica en el paradigma interpretativo, pues su intención es comprender profundamente las vivencias individuales sobre la evaluación y los significados que las personas investigadas producen (Landi, 2023). La metodología cualitativa usada se concreta en relatos biográficos, que permiten comprender directamente las voces de los y las

estudiantes, dándoles la oportunidad de expresar libremente sus sentimientos e interpretaciones de las experiencias educativas (Bolívar, 2002; León, 2015). Esta técnica se ubica en la investigación narrativa y tiene como base la hermenéutica y la fenomenología, corrientes que enfatizan la interpretación y el significado personal para entender las experiencias humanas (León, 2015). Este enfoque permite captar cómo los y las estudiantes interpretan y dan sentido a situaciones específicas de evaluación, conectándolas con sus experiencias previas (Mateos y Núñez, 2011).

En el ámbito de la educación, especialmente en EF, la investigación narrativa ha demostrado ser una herramienta versátil que se adapta a diversos contextos y facilita una amplia variedad de expresiones académicas y creativas. Los datos recogidos mediante esta metodología han sido utilizados para presentar resultados de manera innovadora, como en poemas, cuentos y canciones, reflejando la complejidad y profundidad de las experiencias educativas (Carless y Douglas, 2011; Devís-Devís, 2017; Dowling et al., 2013; Groves y Laws, 2003; Martos y García, 2023).

5.1.1. Participantes

En la presente investigación participaron 60 estudiantes de la Facultad de Educación de una universidad de la Región del Maule en Chile, pertenecientes a 6 grados diferentes: Pedagogía en Lengua Castellana y Comunicación (LC), Pedagogía en Inglés (PI), Pedagogía en Educación Física (PEF), Pedagogía en Educación Parvularia (PEP), Pedagogía General Básica (PGB) y Pedagogía en Historia y Geografía (PHG). En el instrumento se consultó por el género, dando las opciones de femenino, masculino y no binario, resultando la distribución que se observa en la Tabla 19.

Tabla 19. Distribución según género de las personas participantes.

Disciplina	Femenino	Masculino	Cantidad
Pedagogía en Lengua Castellana y Comunicación	7	3	10
Pedagogía en Educación Física	3	7	10
Pedagogía en Inglés	5	5	10
Pedagogía en Educación Básica	7	3	10
Pedagogía en Historia y Geografía	3	7	10
Pedagogía en Educación Parvularia	10	0	10
Total	35 (58%)	25 (42%)	60 (100%)

Posteriormente, se invitó al alumnado a participar en dos grupos focales, invitación que aceptaron 18 estudiantes, Puede observarse la distribución por género en la Tabla 20:

Tabla 20. Participación Grupos focales.

Grupo focal	1	2	Total
	N (%)	N (%)	N (%)
Género			
Femenino	3 (37,5)	8 (80)	11 (61)
Masculino	5 (62,5)	2 (20)	7 (39)
Disciplina			
P. Educación Física	3 (37,5)	4 (40)	7 (39)
P. En Educación Básica	4 (50)	4 (40)	8 (44)
P. Historia y Geografía	1 (12,5)	1 (10)	2 (11)
P. Inglés	0 (0)	1 (10)	1 (6)

5.1.2. Procedimiento

La recogida de datos inicial se llevó a cabo mediante la técnica narrativa de los relatos biográficos para, siguiendo a Argüello (2014), analizar un aspecto pedagógico como es la evaluación, a partir de las experiencias individuales de quienes fueron estudiantes. Al igual que hicieron Valencia-Peris y Lisandra (2018), se pidió al alumnado participante que redactara sus recuerdos acerca de la evaluación en EF. En esta investigación, las personas participantes fueron organizados en grupos de diez y conducidos a una sala de computación, donde cada uno contó con un equipo individual para completar el formulario. Este instrumento incluía, en su primera sección, preguntas de identificación, seguidas de un encabezado que invitaba a relatar libremente sus experiencias en evaluación dentro de la asignatura de EF. Se enfatizó la posibilidad de narrar los recuerdos en el estilo que cada participante considerara adecuado, incluyendo no solo hechos concretos, sino también las sensaciones y emociones asociadas a sus vivencias. El tiempo destinado para la recolección de estos relatos narrativos osciló entre los 45 minutos y 1 hora.

Este procedimiento contó con la aprobación del comité de ética de la Universidad participante (Acta N.º CEC 38-21), además de la autorización correspondiente de la Facultad de Educación de la Universidad de Autónoma de Chile. Las 60 personas participantes firmaron un consentimiento previa información, por parte del primer investigador, de los compromisos éticos adoptados. Estos incluían el respeto a la dignidad de los y las

participantes, la garantía de confidencialidad y anonimato en el tratamiento de los datos, y la voluntariedad de la participación, asegurando que cualquier persona podía retirarse del estudio sin repercusiones. Asimismo, se garantizó que la investigación se ajustaba a los estándares éticos internacionales, como la Declaración Universal sobre Bioética y Derechos Humanos (UNESCO, 2005) y la Declaración de Helsinki, asegurando la protección de los derechos de los y las participantes y la integridad del proceso investigativo.

Una vez recogidos los datos autobiográficos, se procedió a un primer análisis (como se explica en el apartado siguiente), lo que informó el diseño posterior de los dos grupos focales, en los que participaron 18 personas. Cada uno de los grupos focales contó con la moderación del investigador principal, además de una ayudante previamente capacitada para tomar notas y manejar la grabadora de voz. Estas conversaciones tuvieron una duración aproximada de 45 minutos y se realizaron de manera presencial en las dependencias de la universidad participante. Posteriormente se procedió a la transcripción manual de los grupos focales llevados a cabo.

5.1.3. Análisis de datos

El proceso de análisis de datos se realizó conforme a las directrices de la teoría fundamentada (Charmaz, 2014). En la etapa inicial, se optó por un enfoque abierto e inductivo para analizar los relatos narrativos (Corbin y Strauss, 2014). Esta primera fase facilitó la identificación preliminar de categorías emergentes. Posteriormente, para profundizar en estas categorías, se organizaron los dos grupos focales. Los hallazgos obtenidos de estos grupos se incorporaron en un nuevo documento basado en las transcripciones. Se llevó a cabo una nueva ronda de codificación abierta con el objetivo de discernir categorías y subcategorías emergentes (Charmaz, 2014). A continuación, se implementó una codificación axial que proporcionó una comprensión más profunda de las experiencias y perspectivas de los y las participantes del estudio. Aunque en este estudio no se llegó a implementar la codificación selectiva para desarrollar una teoría completa, el uso de las etapas iniciales de la teoría fundamentada permitió una organización efectiva y un análisis preliminar significativo de los datos. Según Bryant y Charmaz (2007), incluso cuando no se completa el proceso entero de la teoría fundamentada, la utilización de sus técnicas iniciales puede aportar claridad y dirección al análisis cualitativo, apoyando la

validez y relevancia de los hallazgos preliminares sin la necesidad de formular una teoría extensa.

Como consecuencia del proceso de análisis, los resultados de esta investigación fueron organizados según la frecuencia de concurrencia identificada en los datos (Tabla 21). Así, dicho análisis muestra que son los principales contenidos curriculares los que aglutinan la mayor concurrencia, siendo estos la condición física, los deportes y las danzas. Estos contenidos curriculares se alinean con el currículum nacional chileno expuesto en las Bases curriculares (Ministerio de Educación, 2015), por lo que se han considerado como categorías pertinentes para organizar la investigación.

Tabla 21. Matriz de categorías emergentes de los relatos narrativos, desde los contenidos curriculares.

Categoría		Ejemplos mencionados	Percepciones negativas	Percepciones positivas
Condición (60)	Física	Test Course Navette (27)	Frustración (10)	Satisfacción (10)
		Abdominales (14)	Desmotivación (8)	Superación (5)
		Correr (14)	Enojo (6)	
Deportes (28)		Flexibilidad (5)	Vergüenza (6)	
		Salto de Cuerda (3)	Malestar físico (6)	
		Test de Cooper (3)	Inseguridad (5)	
		Salto Largo (2)	Desgaste físico (3)	
		Burpees (2)	Inferioridad (3)	
		Circuitos (2)		
		Deporte genérico (12)	Frustración (8)	Satisfacción (6)
		Gimnasia (10)	Discriminación género (4)	Compañerismo (4)
		Fútbol (8)	Desmotivación (3)	Diversión (3)
		Voleibol (7)	Inseguridad (3)	Logro (3)
Danza /Baile (24)		Básquetbol (7)	Discriminación general (3)	
		Handball (5)	Incomodidad (3)	
		Atletismo (2)	Miedo (2)	
		Natación (2)	Ansiedad (2)	
			Vergüenza (2)	
			Injusticia (2)	
		Cueca (10)	Obligación (6)	Entretención (5)
		Bailes tradicionales (10)	Estresante (4)	Satisfacción (3)
		Coreografías variadas (3)	Incomodidad (4)	
		Danzas de Rapa Nui (2)	Monotonía (3)	

El número () corresponde a la frecuencia de concurrencia de los 60 relatos.

Una vez establecidas las categorías de análisis de los datos de los relatos, se procedió al análisis pormenorizado de los grupos focales, lo que añadió tres categorías de análisis que se presentan en la Tabla 22.

Tabla 22. Matriz de categorías emergentes desde los Grupos Focales.

Categorías	Subcategorías	Códigos emergentes
Metodología	Estilos de clase	Enfoque deportivo y realidad de juego Acondicionamiento físico Tiempo libre en hora de clases En edad infantil, gran relevancia al juego
	Planificación	Escasa claridad o ausencia de los objetivos Cambios de temas de forma brusca Falta de continuidad en las unidades
Evaluación	Propósitos	Solo para calificar Como recompensa por competir o participar Sin explicar el sentido Por cumplir Por medir la condición física Fines formativos Retroalimentación en base al disfrute de la clase
	Instrumentos	Rúbricas Observación (sin registro) Lista de cotejo Tablas de rendimiento físico Completa ausencia de instrumentos o indicadores
Actitudes	Injusticia social	Comentarios de burla Estereotipos de género en deportes Comentarios de discriminación por sobrepeso u obesidad Sentimiento de discriminación por competencia motriz
	Deseo de deserción	Por sentirse torpe Por injusticia Por incomodidad (por no rendir) Monotonía Escasa libertad Excusas con apoyo de padres

Elaboración propia en base a los grupos focales

Para la presentación de los resultados, y como alternativa al formato académico tradicional, se utilizaron técnicas literarias para introducir citas textuales del estudiantado. Las citas ayudan a exponer interpretaciones y reflejar los resultados. Para localizar dichas citas se utilizó un número de identificación de cada informante, por ejemplo, 2HGF, siendo HG Historia y Geografía (por los estudios que cursaba) y la F correspondiente al género femenino. En el caso de los grupos focales, se utilizaron pseudónimos, señalando el número del grupo focal evocado (ejemplo: GF2 Valentina; que corresponde a grupo focal 2, con la participante de pseudónimo Valentina).

5.2. Resultados

Los datos recogidos en los relatos narrativos fueron contundentes en el reconocimiento de los test físicos como elemento primordial en la evaluación de la condición física, aspecto al que hizo referencia la totalidad de los informantes. Algunas de las pruebas mencionadas fueron la Course Navette, los abdominales, test de flexibilidad y el hecho de correr y dar vueltas por la cancha, por mencionar algunos. En este sentido, los datos de los grupos focales lo aclaran de la siguiente forma: “las evaluaciones en la asignatura eran siempre las mismas, las famosas pruebas por correr o por hacer ejercicios, que aumentaban acorde pasábamos de curso, cada año se exigía más”- 10PI F. Además, se apuntó al uso de escalas de valoración diferentes en función del género del alumnado: “El profesor tenía tablas de rendimiento para hombres y mujeres, claro que a los hombres le exigía más”-GF2 Claudio.

Además, según los datos recabados, las pruebas aplicadas para evaluar la condición física lo hacían, muchas veces, sin una planificación o secuencia clara. Tal como se observa en la Tabla 22, dentro de la subcategoría "Planificación", los participantes señalaron que, si bien el contenido curricular se aplicaba mediante circuitos, rutinas y test de condición física, estos carecían de una explicación o propósito pedagógico claro. Al respecto, una participante expresó: "Se nos tomaba todos los años el famoso test de naveta, pero nunca se nos explicó lo que medía o reflejaba, tampoco se usaba para nada, solo para poner una nota..."- GF1, Carolina.

Percepciones positivas asociadas a la condición física

En los relatos narrativos se mencionaron emociones positivas asociadas al tipo de evaluación desplegada en el contenido de la condición física, como el gusto y la satisfacción por cumplir con las expectativas del profesorado. De la misma forma, en los grupos focales se hizo hincapié en la satisfacción sentida por aquellas personas que mostraban un buen desempeño físico en las pruebas ejecutadas. Llama la atención como algunas personas explicitaron el gusto por esforzarse para cumplir las expectativas, poniendo encima de la mesa el espíritu de superación. Alguna cita así lo atestigua: “Sí, yo era de las buenas para el deporte, me gustaba correr y siempre sacaba 7,0 [nota máxima], por lo que me motivaba a seguir entrenando” -Laura GF1.

Percepciones negativas asociadas a la condición física

Aunque existen participantes satisfechos con la evaluación experimentada cuando se evaluaba la condición física, la mayoría de los testimonios recogidos destacan aspectos negativos relacionados con estas vivencias. Concretamente, surgieron emociones como frustración y sentimiento de injusticia al utilizar como único criterio de evaluación el rendimiento físico para asignar una calificación:

“Siempre fue una gran batalla responder de forma correcta a todas las actividades evaluativas de Educación Física, porque siempre fui la más pequeña y a todos nos exigían el mismo rendimiento. Siempre fue frustrante ya que, para mí, las calificaciones eran importantes y nunca lograba la exigencia que se me pedía” -7 EP F.

“[...] entonces que solo te califiquen por cuanto corres, es algo frustrante, ya que las notas eran super importantes para mí, para quedar en la universidad, las becas” GF2 Luciano.

En esta dirección, algunas de las quejas realizadas por el estudiantado están relacionadas con la falta de una explicación pertinente sobre la idoneidad de usar dichos test o su relación con la salud:

“Ahora que estudio EF, pude saber la relación que tiene el test naveta con el VO2MÁX. y su relevancia en la salud, pero en el colegio nunca explicaron nada...” -GF1 Cristian.

Los relatos narrativos pusieron de manifiesto, en múltiples ocasiones, sensaciones desagradables y de frustración (ver Tabla 3, percepciones negativas), sobre lo que se profundizó en los grupos focales aludiendo al deseo de abandonar la clase de EF:

“Sí, yo misma pedía a mis papás justificativos para no hacer EF, ellos me apoyaban.”
– GF2 Claudia.

“Odiaba esos test de condición física, me esforzaba hasta que me dolía todo el cuerpo al día siguiente” 5PH M.

En los grupos focales, los y las participantes recordaron estas experiencias y comentaron que eran comunes las excusas relacionadas con el periodo menstrual en mujeres o justificativos firmados por apoderados para evitar la clase de EF. También mencionaron que,

en ocasiones, cuando la clase de EF era en la primera hora de la jornada, optaban por llegar más tarde a las otras clases.

Contenido curricular: los Deportes

Las experiencias de evaluación en el contenido curricular de deportes aparecieron en 28 de los 60 relatos narrativos. En 12 ocasiones se hizo una mención genérica, mientras que en muchas otras se habló de deportes concretos: Gimnasia (10), Fútbol (8), Voleibol (7), Básquetbol (7), Handbol (5) y Natación (2). El Atletismo se menciona en 2 ocasiones, aunque en otras 14 se hizo mención del hecho de correr como procedimiento evaluativo:

“Las evaluaciones de la asignatura de Educación física constaban en su mayoría de aprender a practicar diferentes tipos de deportes como por ejemplo básquetbol, voleibol, entre otros. Este tipo de evaluaciones eran de poco significado para mí...”- 4EP F

Dentro de los procedimientos evaluativos mencionados destacan dos: a) los fundamentos técnicos de cada deporte, de forma aislada, y b) la realidad de juego. En algunos otros casos se mencionaron evaluaciones teóricas relacionadas con los deportes, mediante pruebas escritas o disertaciones sobre las generalidades o reglas del deporte:

“La evaluación consistía en hacer un saque bajo de voleibol y pasar la malla, si la pasábamos era una buena calificación (con factores extra de técnica), en el caso contrario, una calificación reprobatoria...”- 10EF F

Los datos explicitaron diferencias en el proceder de las pruebas de evaluación con un mismo grupo, ya que las alumnas debían rendir pruebas asociadas a deportes como el voleibol y la gimnasia, mientras que sus compañeros varones vivenciaban deportes diferentes como el fútbol o básquetbol. Al profundizar al respecto, se evidenció que hay centros educativos que dividen las clases según el género, con docentes y contenidos diferentes: “En mi colegio era típico que en EF se les pase una pelota a los hombres para jugar futbol y nosotras las mujeres nos fuéramos a bailar o hacer otros juegos en algún espacio disponible” – GF1 Francisca. Por un lado, esto producía un sentimiento de injusticia entre las alumnas, puesto que consideraban no tener las mismas oportunidades que los alumnos. Sin embargo, hubo comentarios por parte de mujeres que valoraron esa división como positiva, ya que con sus

compañeras se sentían a gusto y en confianza, evitando situaciones bruscas o de riesgo con sus compañeros. En los relatos también hubo comentarios negativos sobre esta práctica por parte de hombres, cuando veían que a sus compañeras les exigían menor rendimiento o las dejaban “apoyar las rodillas en las flexiones de codo” – GF1 Juan.

Percepciones positivas asociadas a los deportes

Los relatos valoraban positivamente las evaluaciones asociadas a la práctica de deportes cuando las y los estudiantes cumplían con los requerimientos establecidos, destacándose en estos casos sentimientos de satisfacción y disfrute. Por otro lado, en el contexto de los deportes colectivos, los datos evidenciaron emociones positivas asociadas al ambiente grato que se generaba entre el estudiantado:

“Fue un momento grato ya que aprendí tácticas del fútbol y ser lateral izquierdo en defensa, la evaluación era por el desempeño que teníamos en el área y me fue muy bien” -10HG M

“Mi colegio tenía un enfoque deportivo, las evaluaciones eran también jugando, lo pasamos muy bien y aprendíamos entre compañeros” -Cristian GF2

Como se ha mencionado anteriormente, los deportes eran usualmente asignados según el género, donde los varones practicaban fútbol, muchas veces sin buscar conseguir ningún objetivo de la asignatura, mientras que las mujeres practicaban gimnasia o danzas con música, también con bastante libertad. Esta libertad generaba sentimientos agradables, como se aprecian en la Tabla 3, en la categoría de deportes, que también indica que las percepciones positivas aparecen con menor frecuencia.

Percepciones negativas asociadas a los deportes

Al igual que ocurre en el caso de la condición física, en el plano negativo es la frustración la que aglutina mayor representatividad, hecho que se confirmó en los grupos focales. Así, algunos ejemplos mencionados en los que el alumnado se sintió frustrado fueron aquellas tareas aisladas en las que el resultado determinaba la calificación. Los lanzamientos a la canasta de básquetbol o los saques en voleibol se usaban como elemento de calificación sin que el alumnado recibiera una explicación clara sobre los indicadores de evaluación:

“Recuerdo que pasé última, no podía hacer el saque, y me dejaba repetir hasta lograrlo, me regañó y cuestionó durante todas mis ejecuciones al no poder lograrlo, tocaron el timbre del recreo y continuaba intentándolo, me tenía muy ansiosa e inquieta (...) finalmente no pude concretar el saque y obtuve una muy baja calificación. Nunca vi la pauta...” (10 EF F)

“En el salto del caballete me sentía muy frustrada, nunca pude hacerlo en clases y en la evaluación me calificaron mal...”-6LC F

Además, de forma específica el miedo y la ansiedad hicieron su aparición en la evaluación de la gimnasia artística, por ejemplo, en el salto al caballete, la invertida o parada de manos y las ruedas de carro, entre otras destrezas de suelo. Al profundizar sobre aquello en los grupos focales, se manifestó que tenía que ver con el sentimiento de inseguridad, al no establecer una secuencia progresiva de enseñanza, obligando al estudiantado a rendir mediante la técnica de ejercicios específicos:

“No me gustaba realizar las pruebas, porque el nivel de exigencia era muy alto, aparte me daba susto hacer volteretas hacia atrás o saltar el caballete y no las realizaba, por ende, me bajaban la nota”-1EP F

En el caso específico de los deportes, surgieron quejas relacionadas con los estereotipos de género, ya que las estudiantes no pudieron acceder a las mismas oportunidades o variedades deportivas que sus compañeros varones:

“Veíamos a nuestros compañeros varones realizar actividades entretenidas de Educación Física, como así son el fútbol, el baloncesto o simplemente la quemada, desarrollando el gusto por la asignatura e incluso tomando evaluaciones acordes a esas actividades. Mientras que nosotras las niñas, teníamos que ser “damas” y nuestras evaluaciones [...] constaban de test de flexibilidad, pararnos de manos en la pared o realizar cientos de volteretas”-8 PB F

Hubo casos en que se mencionaron algunas microagresiones por razones de género, como comentarios del tipo “no seas una niñita”, refiriéndose a situaciones de miedo o de bajo rendimiento. Sumado a esto, hubo dos casos que manifestaron incomodidad por parte del profesorado, ya que para ellas era común sentirse atacadas por el docente, no de forma física,

pero si con comentarios fuera de lugar. “Era común que nos separaran en las clases según el deporte o actividad, además de los típicos comentarios de “no corras como niñita”, por parte de compañeros e incluso del profesor...” Carla GF1.

Además de esto, el término “gordofobia” fue utilizado en los relatos por más de una persona, refiriéndose a las burlas y situaciones de discriminación que se vivenciaban por parte del estudiantado o inclusive del docente, con comentarios o muecas que denotaban incomodidad y burlas para los estudiantes con sobrepeso u obesidad:

“No me pasó a mí, pero sí pude ver como sufrían mis compañeros con sobrepeso u obesidad, por burlas de compañeros o incluso algunos comentarios del mismo profesor, para referirse de mala forma a los gorditos” -GF2 Ester.

Contenido curricular: las Danzas

Las danzas tradicionales se rebelaron como un contenido recurrente durante las fiestas patrias de Chile, específicamente en el mes de septiembre. En este sentido, el baile más presente fue el de la cueca. Los procedimientos de evaluación de las danzas se resumen en la representación de dicha danza, donde la mera participación, era conducente a la calificación máxima. De hecho, en algunos relatos se mencionó que se esperaba esta época del año para subir el promedio de calificaciones, ya que se le otorgaba gran realce a la participación de los actos de fiestas patrias.

Percepciones positivas asociadas a las Danzas

Las menciones con una connotación positiva tenían una clara asociación al disfrute y al entretenimiento que el alumnado sentía durante la preparación de las coreografías, en parte por la posibilidad de poder crear y unir los pasos para generar de forma colaborativa las coreografías, que culminaban en representaciones:

“Los bailes de diversos géneros me gustaban mucho, ya que se realizaban en grupo por afinidad, donde luego se debía exponer en frente de todo el curso con una coreografía que nosotras debíamos inventar, esto permitía que con mis amigas nos pudiéramos de acuerdo en los pasos que debíamos realizar”-4 PB F.

“En septiembre podía subir el promedio, donde se realizaban los típicos bailes que cambiaban cada año, en una presentación frente a público... siempre era un 7,0”-8 EF M.

Así, el valor social de estas actividades se destacó positivamente, también por la facilidad que aporta el danzar en grupo a la hora de superar la vergüenza.

“Una evaluación conocida y que se repitió durante enseñanza media y básica, fue el de las fiestas patrias, el típico baile que no variaba de la cueca, la guaracha y el Sau Sau. No me molestaba, creo que bailar era entretenido, presentarnos en público vestidos con trajes típicos me ayudo a perder la vergüenza y subir el promedio de Educación Física, ya que era un 7,0 por bailar”-8PB F

Los datos ponen en valor, además, la posibilidad de ser evaluado o evaluada de otra forma que no fuese bailando, pues algunos comentarios apuntaron a la posibilidad de participar en la confección de vestimentas o en la elaboración de las escenografías de los actos.

Percepciones negativas asociadas a las Danzas

En lo que respecta a las emociones y sentimientos negativos, los relatos apuntan a la monotonía y la obligatoriedad por cumplir, año tras año, instancias de evaluación y enseñanza muy similares, donde el alumnado que se negaba a participar era amenazado con la calificación mínima. Los procedimientos de evaluación relacionados con la danza generaban situaciones de estrés para quienes no se sentían en condiciones de expresarse en público en un acto con bastante presencia de personas. Esta incomodidad venía reforzada por el hecho de sentirse obligado a participar de un procedimiento de evaluación poco claro, ya que no se presentaban los indicadores correspondientes:

“... otras de las evaluaciones más comunes cuando se acercaba la fecha del “18 de septiembre” los típicos bailes como la cueca, costillar, etc. Para mi gusto, era una imposición que más que nada se cumplía por la nota, sin una mayor motivación o explicación de fondo...”-10PB F

“Cuando estuve en la educación básica nunca tuve evaluaciones más allá de bailar cueca, por lo demás, siempre nos ponían 7.0. Esta experiencia no me ayudó a sentirme atraída por la actividad física...”-3PI F

Sobre los procedimientos evaluativos percibidos por el estudiantado

Para empezar, merece la pena resaltar que los datos analizados ponen de manifiesto como una de las críticas recurrentes vertidas es la falta de propósito de las pruebas de evaluación, hecho constatado tanto en los relatos como en los grupos focales. Así, la ausencia de una explicación o el trasfondo de los procedimientos evaluativos realizados, por ejemplo, el test course navette, fue habitual:

“todos los años rendíamos el test course navette, pero nunca supe para qué ni qué relación tenía con la salud...” -GF2 Juan.

“Mi profesor nos decía que si teníamos buen comportamiento nos daba una clase libre, para practicar el deporte que quisiéramos, y cuando había que evaluar era para colocar una nota en el libro, no veía otro fin” -GF1 Luis.

En el caso de haber un propósito, este se relacionaba con el rendimiento físico y deportivo en pruebas estandarizadas para todo el estudiantado. En estos casos, se mencionó en repetidas ocasiones que el propósito era simplemente calificar, para cumplir con la cantidad de notas solicitadas para el año escolar, es decir, como un procedimiento para dar respuesta a un requerimiento administrativo y no para fomentar el aprendizaje:

“En las evaluaciones el profesor solo veía si completábamos las vueltas que nos pedía a una cancha, sin explicar los beneficios, solo por cumplir con la cantidad de notas...”2PI M

De hecho, los argumentos expuestos cuestionan de forma contundente la falta de indicadores claros a la hora de evaluar, ya que simplemente se asignaba una calificación solo por el hecho de participar en las clases o bien representar al establecimiento en alguna competición deportiva. Es decir, se utilizaba la evaluación como un premio o recompensa por participar:

“En los juegos predeportivos nos dividía a hombres de mujeres, nunca nos explicaban los indicadores de evaluación o lo que pretendía evaluar al momento de ejecutarlo...”- 5EF M

“Era evidente el “ojímetro” del profesor, que al final calificaba de una forma que solo el entendía, pero como las notas no eran bajas, nadie decía nada” GF1 Gabriel.

Algunos participantes señalaron que, en algunos casos, el sentido era formativo y el profesorado si explicaba el trasfondo de la evaluación, generando instancias de evaluaciones de proceso, retroalimentación y denotando una planificación de la unidad, lo que se valoraban positivamente:

“Yo recuerdo haber tenido un buen profesor, que demostraba una secuencia en sus objetivos, nos explicaba las evaluaciones, teníamos autoevaluaciones, y buenas retroalimentaciones, en mi caso no puedo estar de acuerdo con el resto...” Francisca GF1.

Cuando se consultó por la retroalimentación en los grupos focales, el alumnado expuso que los y las docentes se preocupaban por consultar sobre el disfrute experimentado en las actividades de clase, incluyendo las de evaluación, con preguntas como ¿Qué fue lo que más les gustó de la clase? o ¿Qué parte de la clase o de la evaluación presentó mayor dificultad para usted?

El último propósito percibido de la evaluación en las experiencias del estudiantado estuvo relacionado con un uso instrumental de esta, en forma de amenaza o castigo. Así, se constata como se utilizaba la calificación mínima (1,0) como instrumento para forzar la participación de la clase:

“...a pesar de que varias mujeres no queríamos participar, este [el docente] nos repetía que si no hacíamos nada nos pondría un 1.0”-9PB F

En sentido contrario, pero igualmente instrumental, la evaluación se usaba como premio por participar en alguna actividad como competencias o representando al establecimiento educativo:

“Recuerdo que nos ofreció un 7,0 en la asignatura si bailábamos en la fiesta folclórica del colegio...”- 5PB F

“Yo competía en atletismo, y cuando representaba al colegio, era un 7,0 directo al libro de calificaciones ...” 10EF F

5.3. *Discusión*

Los hallazgos de esta investigación reflejan que la evaluación en EF sigue estando anclada en un paradigma tradicional, donde las pruebas físicas y de habilidad deportiva son predominantes, replicando un enfoque históricamente instaurado en la disciplina (Hay y Penney, 2009; Oliver-Álvarez y Martos-García, 2023; Portes-Júnior et al., 2022; Thorburn y MacDonald, 2011). Se evidencia que los test de condición física son los instrumentos de evaluación más utilizados, aplicados sin una planificación clara y sin un propósito pedagógico explícito, lo que refuerza una visión reduccionista de la evaluación en EF (López et al., 2006). Además, se identificó el uso de pruebas deportivas de precisión, como los lanzamientos a la canasta en baloncesto, en donde el éxito o la calificación dependía exclusivamente de la cantidad de aciertos, sin considerar el proceso de aprendizaje del estudiantado.

Desde una perspectiva educativa, esta forma de evaluar contrasta con principios clave del aprendizaje significativo y auténtico, los cuales sugieren que la enseñanza debe tener un sentido claro para el estudiantado (Ausubel et al., 1978; Lombardi, 2007). La falta de explicaciones sobre los procedimientos evaluativos generó en el alumnado cuestionamientos respecto a los propósitos de la asignatura, lo que va en contra de la idea de que la motricidad debe ser comprendida y experimentada con un sentido claro y orientado al bienestar (Sérgio, 2006).

Por otro lado, si bien algunos participantes expresaron satisfacción al cumplir con las expectativas en las pruebas de condición física o en el rendimiento deportivo (normalmente en el caso de alumnado que cuenta con una buena competencia motriz), en la mayoría de los casos se evidenciaron emociones negativas, como frustración e injusticia, especialmente al considerar que las calificaciones tienen un impacto en el acceso a la educación superior y otros beneficios académicos (Danthony et al., 2020; Duque et al., 2021; Oliver-Álvarez y Martos-García, 2023). Estas emociones, asociadas frecuentemente a prácticas evaluativas

orientadas exclusivamente desde la ideología del rendimiento (Molina y Beltrán, 2007), coinciden con estudios que han documentado cómo los recuerdos escolares de evaluación, tanto en el profesorado en formación como en ejercicio, pueden generar ansiedad, inseguridad o desmotivación, afectando su posterior proyección profesional (Atienza et al., 2018; Saiz-González et al., 2025; Sonlleve et al., 2018). Además, se observó que las evaluaciones eran utilizadas como elementos de recompensa o castigo, lo que refuerza prácticas descontextualizadas del aprendizaje y centradas en la regulación externa del comportamiento estudiantil (Cañadas et al., 2021; López et al., 2006;).

Las experiencias evaluativas en los deportes reflejaron una tendencia similar, con un énfasis en la ejecución técnica aislada y una menor valoración de la realidad de juego. Las diferencias de género en la asignación de deportes y las oportunidades de aprendizaje fueron aspectos relevantes en los relatos, destacándose la perpetuación de estereotipos de género en la enseñanza de la EF (Alvariñas-Villaverde y Pazos-González, 2018; Pastor-Vicedo et al., 2019; Piedra et al., 2013). Mientras algunas estudiantes valoraban positivamente la separación por género en las clases de EF porque les generaba confianza, otras la percibían como una limitación a sus oportunidades de aprendizaje, lo que se alinea con investigaciones que destacan la necesidad de transformar la cultura de la asignatura hacia una perspectiva de igualdad e inclusión (Lleixà et al., 2020; Matus-Castillo et al., 2023).

Además, emergieron experiencias de discriminación por género expresadas a través de comentarios sexistas tanto del profesorado como de los compañeros, lo que contribuyó a la desmotivación de las estudiantes y al abandono de la clase (López et al., 2022; Monforte y Úbeda, 2019). Este tipo de comentarios y actitudes han sido documentados en la literatura como factores que alejan a las niñas y mujeres de la actividad física y deportiva, perpetuando desigualdades estructurales en la EF (Scharagrodsky, 2004).

Asimismo, la "gordofobia" apareció como una experiencia recurrente en los relatos y grupos focales, manifestándose en la discriminación hacia estudiantes con sobrepeso u obesidad, tanto por parte de sus pares como del profesorado. Este fenómeno ha sido documentado como un factor de exclusión en la EF, asociado a la estigmatización del cuerpo y la presión social en torno a la imagen corporal (Martínez y Bastarrica, 2023). Investigaciones previas han demostrado que esta discriminación puede provocar el abandono de la asignatura y la evitación de la actividad física, afectando el bienestar emocional y físico

del estudiantado (da Costa et al., 2024; Devís-Coret y Martos-García, 2023a). En este sentido, enfoques como las pedagogías de la corporeidad han sido destacadas por su potencial para fomentar la reflexión crítica, el reconocimiento de la diversidad corporal y la construcción de experiencias más inclusivas en la clase de EF (Aartun et al., 2020).

Otro hallazgo relevante fue el uso instrumental de la evaluación, tanto como castigo para quienes no participaban en las actividades como recompensa para quienes representaban al colegio en competencias deportivas (Quirós, 2021). La literatura ha señalado que estas prácticas no solo desvirtúan el verdadero sentido de la evaluación, sino que refuerzan dinámicas de exclusión y presión social dentro de la asignatura de EF (Cañadas et al., 2021; López et al., 2006).

En cuanto a la evaluación de las danzas como contenido curricular, los datos indicaron que esta se realizaba principalmente en el contexto de celebraciones patrias, con un enfoque centrado en la mera participación, sin criterios claros de evaluación. Esta imposición generaba desmotivación en quienes no se sentían cómodos con la exposición pública o que no encontraban un sentido formativo en la actividad. En este sentido la evidencia señala que, cuando la enseñanza de la danza en EF se basa en un paradigma participativo y con una planificación consciente, puede favorecer el bienestar social, la promoción de la cultura y la actividad física (Adolfo et al., 2023; Smith y Pérez, 2023).

Finalmente, la falta de claridad en los propósitos de la evaluación emergió como un elemento central en los relatos del estudiantado. La ausencia de una explicación sobre los procedimientos evaluativos y su relación con la salud y el bienestar generó confusión y desmotivación. En contraste, los pocos casos en que se aplicó una evaluación formativa y compartida, con instancias de retroalimentación y autoevaluación, fueron valorados positivamente por los y las estudiantes (Cañadas et al., 2021; Leenknecht et al., 2021). Estos resultados coinciden con investigaciones que han señalado que cuando los recuerdos escolares de evaluación integran experiencias más inclusivas, dialógicas y reflexivas, estas pueden potenciar la motivación hacia el aprendizaje y favorecer el desarrollo de prácticas evaluativas más humanas y significativas en el futuro desempeño docente (Hortigüela et al., 2016; Muñoz-Sepúlveda et al., 2025). Se refuerza la necesidad de seguir investigando sobre la EF escolar (Moreno y Poblete, 2015).

5.4. Conclusiones estudio 2

Los hallazgos de esta investigación revelan percepciones ambivalentes respecto a la evaluación en la asignatura de EF, lo que refuerza la necesidad de revisar y transformar las prácticas evaluativas. Aunque los contenidos evaluados (condición física, deportes y danzas) se ajustan al currículo nacional, las estrategias predominantes responden a un enfoque tradicional centrado en el rendimiento físico y técnico, lo que afecta negativamente a la motivación y la adherencia a la actividad física, especialmente cuando las evaluaciones carecen de propósito claro y retroalimentación formativa.

Algunos estudiantes expresan satisfacción al cumplir con los criterios, pero una parte significativa experimenta frustración y desmotivación al no alcanzar los estándares exigidos. La falta de claridad y criterios inclusivos agrava este malestar, generando estrategias de evitación como la inasistencia o la justificación médica. Además, el estudio visibiliza prácticas discriminatorias relacionadas con el género y el estatus de peso, que refuerzan estereotipos y afectan al bienestar emocional del estudiantado.

Una actitud general de desmotivación se vincula directamente con un enfoque evaluativo que no promueve el aprendizaje significativo. Avanzar hacia una evaluación formativa y compartida, centrada en la participación activa y en la retroalimentación constructiva, puede fortalecer la motivación y mejorar la experiencia en EF.

En conclusión, esta investigación destaca la importancia de repensar las prácticas evaluativas en EF, adoptando enfoques más inclusivos y amplios que consideren las necesidades y experiencias del estudiantado. Solo así se conseguirá promover un entorno educativo más equitativo y efectivo, que verdaderamente fomente el desarrollo integral de todo el estudiantado.



VNIVERSITAT DE VALÈNCIA

CAPÍTULO 6

ESTUDIO 3

Estudio 3. Del rendimiento al esfuerzo y actitud: Enfoques evaluativos en Educación Física desde la experiencia docente.

Estado: Artículo en revisión.

Resumen

El objetivo de esta investigación fue comprender cómo las experiencias evaluativas vividas durante la etapa escolar, junto con la formación inicial y la propia práctica docente, han influido en la adopción de enfoques evaluativos por parte de profesorado chileno. Se llevaron a cabo entrevistas semiestructuradas y análisis de instrumentos de evaluación de diez docentes de Educación Física en ejercicio. Los resultados revelan cambios en las prácticas evaluativas que van desde experiencias escolares con métodos tradicionales centrados exclusivamente en la medición del rendimiento físico hacia enfoques que priorizan el proceso de aprendizaje, el esfuerzo y la actitud del alumnado. Este cambio comenzó en la formación inicial docente y se afianzó durante la etapa de práctica profesional. La investigación concluye que la formación docente y la etapa profesional son fundamentales para la transformación de la praxis docente hacia enfoques de evaluación más participativos y equitativos en el ámbito de la Educación Física.

6.1. Metodología

Este estudio cualitativo se enmarca dentro de un enfoque interpretativo, ya que busca comprender las percepciones y experiencias subjetivas del profesorado respecto a la evaluación en EFYS. Según Creswell y Poth (2018), la investigación cualitativa permite estudiar fenómenos complejos desde la perspectiva de quienes los viven, otorgando profundidad y significado a los datos obtenidos.

6.1.1. Participantes

Los participantes fueron seleccionados mediante un muestreo intencional (Palinkas et al., 2015), el cual permite elegir sujetos que tienen una experiencia significativa y pertinente sobre el fenómeno de interés, asegurando así la relevancia de los datos recogidos. Las personas participantes fueron 3 profesoras y 7 profesores de EFYS de la etapa de Primaria

que, en el momento de la entrevista, se encontraban ejerciendo en diferentes establecimientos educativos de la región del Maule, Chile (Tabla 23). Dentro de los criterios de inclusión se consideró, como principal requisito, llevar al menos un año ejerciendo como docente de la asignatura de EFYS en cualquier contexto escolar.

Tabla 23. Caracterización del profesorado entrevistado.

Participante	Género	Rango edad (años)	Años de experiencia docente	Dependencia del centro educativo
1	Mujer	50 a 59	32	Particular
2	Mujer	40 a 49	24	Municipal
3	Hombre	30 a 39	11	Particular-subvencionado
4	Hombre	30 a 39	10	Municipal
5	Hombre	30 a 39	10	Particular
6	Hombre	30 a 39	8	Particular-subvencionado
7	Hombre	30 a 39	7	Particular-subvencionado
8	Hombre	30 a 39	4	Municipal
9	Hombre	20 a 29	4	Municipal
10	Mujer	20 a 29	2	Municipal

Elaboración propia

6.1.2. Técnicas de recogida de datos

6.1.2.1. Entrevistas semiestructuradas

Las entrevistas semiestructuradas permitieron explorar las experiencias personales y las conceptualizaciones de los y las participantes de manera flexible (Hernández, 2014; Rubin y Rubin, 2020). Tuvieron un guion con cuatro secciones principales: (1) una introducción orientada a la caracterización del contexto de los entrevistados, (2) una exploración sobre su conceptualización de la evaluación, (3) un análisis de sus experiencias evaluativas en la etapa escolar y durante la formación universitaria, y (4) una indagación sobre sus prácticas evaluativas actuales en el ejercicio profesional.

El diseño del guion de la entrevista semiestructurada se construyó siguiendo un enfoque sistemático y fundamentado (Troncoso y Daniele, 2014). Una vez diseñado en base a la literatura consultada, el guion fue sometido a un proceso de análisis de calidad mediante juicio de expertos. Para ello, se consultó a especialistas en el área, quienes evaluaron la pertinencia, claridad y capacidad de las preguntas para captar información relevante y consistente con los objetivos de la investigación. Las observaciones y recomendaciones de los expertos fueron incorporadas para garantizar la calidad y adecuación del instrumento.

6.1.2.2. Análisis documental

Se solicitó a los y las participantes que facilitaran un instrumento de evaluación que hubiesen utilizado en el semestre académico en curso (tabla 24). De los diez participantes entrevistados, solo cuatro accedieron a enviar sus instrumentos de evaluación mediante un correo electrónico en formato digital para su posterior análisis en conjunto con las entrevistas, con el fin de profundizar en sus concepciones y prácticas. Estos incluyen una rúbrica, dos escalas de apreciación y un registro diario, dirigidos a los cursos de 1° y 2° básico. Los objetivos de los instrumentos varían entre evaluar habilidades motrices básicas (como la manipulación y la combinación de locomoción y manipulación en circuitos), y hábitos de higiene, conducta, participación y esfuerzo. La cantidad de indicadores para cada instrumento oscila entre 5 y 6, evidenciando un enfoque detallado en la evaluación de aspectos específicos relacionados con las habilidades y actitudes de los estudiantes.

Tabla 24. Caracterización de los instrumentos de evaluación analizados.

Participante	Tipo de instrumento	Nivel educativo	Objetivo del instrumento	Cantidad de indicadores
3	Rúbrica	1° básico (primaria)	Evaluar habilidades motrices básicas, especialmente manipulación, a través de circuitos.	6
6	Escala apreciación	2°básico (primaria)	Evaluar habilidades motrices de locomoción y manipulación promoviendo la participación activa.	6
7	Escala apreciación	2°básico (primaria)	Evaluar el combinar habilidades motrices de locomoción y manipulación en circuitos, con buena actitud.	5
10	Registro diario	1°básico (primaria)	Evaluar hábitos de higiene, conducta, participación y esfuerzo,	4

Elaboración propia

6.1.3. Cuestiones éticas y de procedimiento

Esta investigación contó con la aprobación del Comité de Ética de la Universidad Autónoma de Chile (código N°CEC 38-21). Para llevar a cabo esta investigación se solicitó la participación voluntaria al profesorado de EFYS. Una vez aceptada su participación mediante la firma de un consentimiento informado se aplicaron las entrevistas semiestructuradas. Estas fueron realizadas de forma presencial y personalizada con cada una

de las personas participantes. Tuvieron una duración aproximada de 45 minutos, siendo registradas con grabadora de voz, para su posterior transcripción y análisis.

6.1.4. Análisis de datos

Para el análisis de los datos cualitativos obtenidos a partir de las entrevistas y los instrumentos de evaluación, se empleó la metodología de análisis categorial inductivo, una técnica ampliamente reconocida en la literatura académica (Hodge y Sharp, 2016; Pérez-Samaniego et al., 2011). Este enfoque analítico permite una exploración exhaustiva de los datos, comenzando con la identificación de unidades de significado relevantes.

El proceso de análisis se desarrolló tomando como punto de partida las dimensiones principales establecidas en el guion de la entrevista (Rodríguez et al., 2005). La codificación y agrupación temática de las unidades de análisis en categorías y subcategorías emergentes siguió procedimientos fundamentados en el análisis temático, como el descrito por Evans (2018). Este enfoque se destaca por su utilidad para identificar patrones y significados en entrevistas semiestructuradas, así como en el análisis de instrumentos de evaluación. Dicho método permitió una comprensión estructurada y profunda de la información recopilada.

Para cada una de las dimensiones propuestas en esta investigación (experiencias previas en evaluación y proceso evaluativo), emergieron diversos códigos derivados tanto de las entrevistas como de los instrumentos de evaluación. Estos códigos fueron agrupados en categorías y subcategorías, que posteriormente se organizaron en una matriz (tabla 25). Esta matriz sintetiza las relaciones entre las dimensiones, categorías, subcategorías, códigos y descripciones de los hallazgos principales.

Para presentar los resultados se ha evocado al bricolaje de posibilidades que entrega la investigación cualitativa (Moral, 2016), incluyendo citas textuales de las entrevistas (utilizando el número de participantes como identificación, ejemplo P3 = participante 3) o indicadores de evaluación de los instrumentos entregados por el profesorado.

Tabla 25. Matriz de categorías y subcategorías emergentes del análisis de las entrevistas e instrumentos de evaluación.

Dimensión	Categoría	Subcategoría	Códigos		
Experiencias previas en Evaluación	Experiencias en edad escolar	Evaluación tradicional	Criterios estrictos (Ent) Rendimiento físico (Ent) Técnica deportiva (Ent)		
		Ausencia evaluación	No reconoce esfuerzo (Ent) Solo procedimental (Ent) No se evaluaba (Ent)		
	Experiencias universitarias	Evaluación tradicional	Rendimiento técnico (Ent) Asignaturas enfocadas en rendimiento (Ent)		
		Reflexión en enfoque	Cuestionamiento a métodos tradicionales (Ent) Consideración a la diversidad (Ent)		
		Enfoque actitudinal y proceso	Enfoque en proceso (Ent) Asignaturas clave en enfoque actitudinal (Ent)		
	Transición de enfoque	Cambios en sus propuestas	De la ausencia a evaluar (Ent) De enfoque tradicional a integral (Ent)		
Proceso evaluativo	Conceptualización	Enfoque formativo	Progreso (Ent+Ins) Proceso (Ent+Ins) Comunicada (Ent) Formativa (Ent) Más que calificar (Ent)		
			Individualizada	Individual (Ent+Ins) Personalizada (Ent+Ins) Intereses estudiante (Ent) Adecuada (Ent) Accesible (Ent)	
		Foco		Cualitativo (Ent) Esfuerzo (Ent+Ins) Comportamiento (Ent+Ins) Progreso (Ent) Conceptual (Ent+Ins) Procedimental (Ent+Ins) Actitudinal (Ent+Ins)	
				Usos	Diagnóstico (Ent) Formativo (Ent) Calificar (Ent+Ins) Seguimiento de actitud (Ent+Ins) Observación clase a clase (Ent+Ins)
					Prácticas evaluativas
			Dimensiones		

Nota. "Ent": Códigos provenientes de las entrevistas. "Inst": Códigos provenientes de los instrumentos de evaluación.

6.2. Resultados

Experiencias previas en evaluación

En las experiencias en edad escolar recordadas por el profesorado participante, la gran mayoría mencionó evaluaciones orientadas al rendimiento físico o ejecución de técnicas deportivas con criterios estrictos: "En la escuela si hacíamos 10 lanzamientos al aro acertados, sacábamos un 7 (calificación máxima), si no, la nota bajaba" (P5). No obstante, hubo un caso en que se manifestó la ausencia de procesos evaluativos de forma total, inclusive remarcando que las calificaciones aparecían "por arte de magia" (P2). Por su parte, los docentes de mayor edad relataron que, en su experiencia, los estudiantes que no cumplían con las pruebas de rendimiento físico recibían calificaciones bajas:

Recuerdo a compañeras que tenían promedio rojo(reprobatorio) en Educación Física, por no rendir por ejemplo en los test de resistencia aeróbica (P2).

Un elemento en común en todos los recuerdos es la escasa consideración de elementos actitudinales: "Nunca se tomaba en cuenta el esfuerzo, solo el resultado" (P10). Respecto de los recuerdos de la etapa universitaria, estos fueron variados, pero también destaca un enfoque de evaluación asociado al rendimiento físico y/o deportivo: "Nos evaluaban haciendo gimnasia artística, si lograbas el salto, aprobabas, si no, no" (P3).

Sin embargo, también hubo asignaturas, y específicamente docentes universitarios, que marcaron un punto de inflexión, al incluir aspectos del proceso, del esfuerzo y la actitud más allá del rendimiento observable. Este tipo de experiencias se manifestaron con más intensidad en el grupo de entrevistados de menor edad:

En esa asignatura fue cuando entendí que no podía evaluar solo el rendimiento; debía mirar el progreso y las características individuales de cada estudiante (P9).

Todos los participantes reconocieron transformaciones en sus concepciones y prácticas evaluativas al contrastar sus experiencias escolares con su desempeño actual. Se evidencia una clara distancia entre cómo fueron evaluados y cómo evalúan hoy: "Totalmente diferente, de partida yo evaluó, en el colegio ni siquiera se evaluaba" (P2). Algunos identificaron este cambio durante su formación universitaria, "Muy distinta a mi experiencia, en la universidad

me di cuenta de aspectos que sí había que considerar” (P9), mientras que, en otros casos, la transición ocurrió en el ejercicio profesional: “Fue trabajando que me di cuenta que lo que viví necesitaba cambios, no somos iguales, hay que tomar en cuenta el esfuerzo” (P3). En todos los casos, se observa un tránsito desde enfoques evaluativos tradicionales centrados en el rendimiento hacia modelos que valoran el proceso, la actitud y el esfuerzo del estudiantado.

La propia práctica docente como como formación

Aunque la formación universitaria influyó en la concepción que los y las participantes tienen de la evaluación, otro contexto importante lo es el desempeño de su labor profesional como docentes. Así, los datos apuntan a que el ejercicio docente en Educación Física motivó un cambio, prestando mayor atención a la interacción con el alumnado y a la necesidad de adaptar las estrategias evaluativas:

Cambiar mi forma de evaluar fue necesario; no podía aplicar lo mismo que viví en el colegio porque dejaba fuera a muchos estudiantes, por lo que cambié al llegar al colegio y darme cuenta de aquello que generaba (P7).

En años anteriores, cuando la prueba SIMCE estaba de moda, también calificábamos en base a test de condición física, pero no nos dio muchos resultados, finalmente porque la nota era muy baja, y al final era poco consecuente con nuestro pensamiento respecto a lo que queríamos lograr, y se aplicó hasta la pandemia... (P6)

De una forma u otra, los datos reflejan una tendencia hacia el desarrollo de propuestas de evaluación más atingentes al enfoque formativo, y alejadas de la concepción técnica y de rendimiento. Con todo, algunos comentarios apuntan a la coexistencia de las dos formas de entender la evaluación:

En el departamento de Educación Física del colegio existen 2 tipos de profesores, los profesores que son más antiguos y los profesores que son nuevos. Los más nuevos trabajamos en todo el tema de evaluar lo procedimental y sobre todo lo actitudinal por sobre el rendimiento físico, pero de repente se nos quedan los profesores que son más antiguos que siguen calificando por capacidad física... (P3)

Hemos tenido diferencias de opinión y de prácticas, ya que en el colegio hay profesores que son más tradicionales, debe ser por su edad, donde sus prácticas son bastante militarizadas y enfocadas al rendimiento... (P8)

El concepto de evaluación

En este sentido, el concepto de evaluación es determinante a la hora de interpretar las diferentes prácticas evaluativas. Así, según el análisis de las entrevistas, en términos generales se asocia la evaluación a un proceso llevado a cabo a lo largo del tiempo y que puede ir asociado una calificación. Según el profesorado, debe ser adaptado a cada estudiante, ya que como menciona P9: “no todos tienen las mismas capacidades y tampoco van a llegar de la misma forma al objetivo”. En esta dirección, parece existir cierto consenso en describir la evaluación como un proceso dinámico y formativo que va más allá de asignar calificaciones numéricas, valora el progreso individual de cada estudiante, y adopta metodologías y criterios de evaluación que garanticen la equidad y el desarrollo integral del estudiantado:

La evaluación como un medio para que el estudiante se motive, más que un medio punitivo al final (P4)

Bueno, la evaluación en general, para nosotros es más que valorar el resultado por una nota, es evaluar el proceso... (P1)

yo considero que la evaluación es un proceso completo, un proceso formativo que no tiene que ver solamente con el resultado final, sino que es un proceso que requiere de un inicio, basado en el interés del estudiante... (P6)

La conceptualización que el profesorado hace de la evaluación incorpora claramente los aspectos actitudinales, tanto en el discurso de las entrevistas como en los instrumentos de evaluación analizados, ya que en todos los casos se consideraban indicadores actitudinales, por ejemplo:

Demuestra entusiasmo y compromiso en todas las actividades (Rúbrica 1º Básico de P3)

Es diferente evaluar que calificar, se debe considerar el proceso, no solo el resultado final, sino como llegó a eso, su participación y actitudes... (P1)

La nota por conducta e higiene corresponde al 50% de la nota final de cada unidad (Registro diario 1 básico de P10)

Prácticas evaluativas

Los participantes identificaron diversos propósitos atribuidos a la evaluación en su práctica docente. Entre ellos, se mencionó su uso diagnóstico al inicio de una unidad o ciclo escolar, ya sea con el objetivo de conocer el estado físico del estudiantado antes de comenzar el año o para evaluar los aprendizajes previos. Asimismo, se destacó la función formativa de la evaluación, utilizada para el seguimiento continuo de actitudes y la participación del estudiantado en clase. No obstante, también se reconoció su dimensión sumativa, vinculada a la necesidad de calificar, dado que la normativa exige la generación de una calificación final al término de cada unidad. En este sentido, la evaluación cumple múltiples roles dentro del proceso de enseñanza-aprendizaje, oscilando entre el seguimiento del desempeño del estudiantado y el cumplimiento de requerimientos institucionales:

Principalmente diagnosticamos condición física, ahora hay que hacer una evaluación diagnóstica por unidad, nosotros tenemos 4 unidades (P2)

La evaluación está presente en todo momento, en el proceso, clase a clase, en las actitudes, aprendizaje, retroalimentación...está en todo momento... (P10)

Respecto a la calificación, en términos generales las notas se generan dos o tres veces al semestre, prácticamente como resultado de cada unidad. Para ello, se obtiene la calificación a través de un instrumento de evaluación que debe estar debidamente validado por la UTP de cada establecimiento. Los instrumentos de evaluación mayormente mencionados por el profesorado fueron las rúbricas, escalas de apreciación y listas de cotejo. Otro instrumento mencionado fueron los registros diarios o “clase a clase” como lo denomina el profesorado, que identifican actitudes y participación del estudiantado en todas las clases de cada unidad didáctica. Si bien el currículum nacional chileno propone instrumentos de evaluación y orientaciones, el profesorado manifestó su preferencia por diseñar sus propios instrumentos, según las características de su contexto, para favorecer el desempeño de su estudiantado:

Preferimos crear nuestros instrumentos, acorde a lo que sea el proceso cada año, y luego se valida por la UTP, no utilizamos los propuestos en las bases curriculares... (Alan)

La apertura hacia la consideración de otras dimensiones, más allá de la puramente procedimental, se ve reflejada, por ejemplo, en la ponderación que se otorga en las calificaciones a los aspectos actitudinales:

...acá en el colegio, todas las calificaciones se ven afectadas por lo actitudinal y su participación clase a clase, se considera con el mismo peso, el 50% de ponderación en la nota de la Unidad... (P3)

El estudiante participa en conjunto con su grupo manteniendo un orden en los turnos (Pauta de evaluación 2° de P7)

La dimensión conceptual de la enseñanza fue la menos señalada dentro del proceso de calificación, de hecho, de los cuatro instrumentos de evaluación analizados, donde se analizaron 21 criterios de evaluación, solo dos de ellos están orientados al desempeño conceptual.

El estudiante reconoce con facilidad la habilidad motriz trabajada (Pauta de evaluación 2° P6).

... por ejemplo, al momento de evaluar manipulación, a los que tienen toda la parte teórica, digamos completa, les adhiero un punto, dependiendo del puntaje total de la evaluación... (P10)

Pero no solo las actitudes o comportamiento han tenido presencia en los datos, sino que los procedimientos han sido señalados como importantes. Así, la focalización de la evaluación en los aspectos procedimentales de la enseñanza trae a colación la valoración del esfuerzo como elemento nuclear, no solo como aspecto a evaluar, sino también como hecho diferencial para la calificación del alumnado:

La calificación se consigue mucho más por el esfuerzo que ellos ponen, por la actitud que ellos colocan más que por la habilidad que ellos puedan tener, porque finalmente todo el currículum de nuestro colegio está planteado para que todos puedan lograrlo... (P6)

Además, la reorientación de las prácticas evaluativas hacia elementos más ligados con las actitudes y el esfuerzo se asocia con dos aspectos importantes: por una parte, el hecho de animar al estudiantado a que participe activamente y muestre respeto por sus compañeros y compañeras, la clase y el profesorado, y, por la otra, a la necesidad de favorecer al estudiantado menos hábil motrizmente:

Yo soy bastante flexible en cuanto a la nota de los niños cuando veo progreso. Pero también soy muy estricta por la nota que tendrán por la conducta, que va incorporada dentro de los ítems. Se les conversa del impacto que tendrá en su calificación, y si no quieren participar o presentan mala actitud, la nota les baja bastante. (P10)

Al final de cada unidad se otorgará un incentivo al esfuerzo a todos los estudiantes que hayan presentado problemas motrices al inicio de la unidad y que se superen respecto de la evaluación diagnóstica realizada al comienzo de cada unidad (Nota de registro diario 1° básico de P10)

En cierta manera, la evaluación se utiliza como un incentivo a la participación del estudiantado, de forma que es la motivación extrínseca la que entra en juego en forma de calificación:

En la unidad de folklore, es un poco más complejo, porque no a todos les gusta bailar. Pero todos tienen que bailar. Y nosotros entendemos eso, entonces evaluamos el proceso, en el cual el estudiante se va poniendo notas, clase a clase por trabajar por participar en el baile, por seguir las instrucciones del profesor(a), por ensayar con su pareja o grupo. (P6)

Los que no tienen notas altas, tiene más relación a conductas que no sean propias de la clase, o que no tenga participación. Pero tenemos estudiantes que se niegan a participar o que están lesionados, etc. nosotros también tenemos trabajo escritos para ellos, para que participen en la asignatura, aunque sea en la biblioteca... (P5)

6.3. Discusión

Los resultados de esta investigación evidencian notorios cambios en las prácticas evaluativas entre las que vivenciaron los y las profesoras de EF chilenos y las que manifestaron aplicar en la actualidad, marcando una transición de un modelo tradicional,

centrado exclusivamente en el rendimiento y la técnica (Castillo-Retamal et al., 2022), hacia enfoques más inclusivos y formativos que valoran tanto el proceso de aprendizaje como el esfuerzo del estudiante (Calatayud, 2019; Martos y García, 2023). Este cambio es congruente con las recomendaciones de Gall (2015), quien sugiere que la evaluación en EF debe centrarse no solo en medir los resultados finales, sino también en proporcionar retroalimentación constructiva que fomente el desarrollo integral del estudiantado (Gallardo et al., 2022). Además, López-Pastor et al. (2012) respaldan esta idea, destacando que la evaluación formativa y compartida permite un proceso de reflexión crítica en los estudiantes, promoviendo su autonomía y responsabilidad en el aprendizaje.

La evaluación tradicional en la etapa escolar del profesorado se caracterizó por el uso de criterios estrictos y cuantitativos, como reflejan los datos. Esto es coherente con lo apuntado por la literatura respecto al carácter hegemónico de este tipo de evaluación. Esta orientación evaluativa que hace un uso predominante de técnicas cuantitativas limita el desarrollo de una comprensión holística del aprendizaje del estudiante en EF (Llapur, 2023), generando una barrera para su progreso (Danthony y Cury, 2020; Oliver-Álvarez y Martos-García, 2023). Asimismo, Foglia et al. (2014) sostienen que las prácticas evaluativas tradicionales no permiten una evaluación continua ni formativa, lo cual es crucial para un desarrollo adecuado en la disciplina de EF (Hortigüela-Alcalá et. al, 2021).

Sin embargo, los datos apuntan a que la formación docente universitaria supuso un punto de inflexión, pues es en esta etapa cuando algunos docentes comenzaron a cuestionar y modificar sus enfoques evaluativos, integrando elementos actitudinales y centrándose en el proceso de aprendizaje. Para ello, se revela fundamental haber tenido experiencias de evaluación en las que se valoraba tanto el proceso como la actitud, como han relatado en las entrevistas algunos docentes. Barba-Martín et al. (2020) enfatizan la importancia de un modelo de evaluación formativa, destacando que este tipo de evaluación favorece la implicación activa de los y las estudiantes y permite romper con las barreras impuestas por los modelos tradicionales (Secchi et. al, 2016; Toro-Suaza y Chaverra, 2022).

Más tarde, inmersos ya en la práctica docente, los y las docentes mostraron su inclinación por incluir en la evaluación indicadores que trasciendan el rendimiento físico o la habilidad motriz, lo cual es coherente con la evidencia en investigación disciplinar de los últimos años

(Atienza et al., 2018; Cañadas, 2020; López-Pastor et al., 2006; Toro-Suaza y Chaverra, 2022). Estos autores señalan que la evaluación en EF debe considerar no solo los conocimientos técnicos, sino también los componentes actitudinales y procedimentales, de forma que se pueda evaluar al estudiante de manera integral y justa (Frossard et al., 2021). Los y las docentes participantes de este estudio destacan que han dejado de lado la evaluación basada únicamente en la ejecución de habilidades técnicas para centrarse en las actitudes y en el compromiso de cada estudiante, lo cual se alinea con la tendencia hacia una evaluación formativa (Bernard y González-Moreno, 2020; Calatayud, 2019; Jiménez, 2017).

En un análisis más detallado, los resultados evidencian que de entre todos los aspectos a evaluar son las actitudes la dimensión que predomina. El comportamiento en clase, la participación y el esfuerzo son los criterios más destacados y valorados por el profesorado, seguidos por el desempeño procedimental y, en menor medida, por los aprendizajes conceptuales. Esta jerarquización coincide con estudios previos que destacan la motivación del estudiantado hacia modelos evaluativos que premian la implicación activa en clase, lo cual suele traducirse en calificaciones altas (Mackley et al., 2022; Sevil-Serrano et al., 2020; Zhu, 2015). Esta tendencia podría evidenciar un desplazamiento progresivo del énfasis en el rendimiento físico o técnico hacia una valoración casi exclusiva de las actitudes, lo cual plantea desafíos en cuanto al equilibrio que favorecería el aprendizaje en esta asignatura.

Esta orientación evaluativa no es exclusiva del contexto chileno. De hecho, se ha documentado una tendencia similar en Latinoamérica, donde, pese a la persistencia de enfoques tradicionales, se observa una incorporación progresiva de criterios actitudinales, con particularidades según el contexto. En Colombia, por ejemplo, se ha llegado a incluir el uso del uniforme como parte de los criterios evaluativos (Toro-Suaza y Chaverra, 2022), mientras que, en Chile, los docentes entrevistados destacaron la observación del comportamiento en clase como un eje transversal. En el caso de España, aunque aún predomina una orientación hacia el rendimiento físico, investigaciones recientes destacan un creciente reconocimiento de enfoques formativos y diagnósticos, donde los aspectos actitudinales también cobran relevancia (Chaverra-Fernández y Hernández-Álvarez, 2019; Zubillaga y Cañadas, 2021). Esta misma lógica ha sido registrada en Estados Unidos, donde

el esfuerzo y la participación han sido ampliamente adoptados como criterios centrales en la práctica evaluativa del profesorado de EF (Killian y Woods, 2021).

En este marco de prácticas evaluativas que privilegian criterios actitudinales, resulta imprescindible considerar la estrecha relación entre evaluación y motivación, especialmente en sistemas donde la calificación incide directamente en el acceso a niveles educativos superiores (Townsend y McNamara, 2021). Si bien la centralidad del esfuerzo y la participación ha sido valorada como una estrategia para fomentar la inclusión y la implicación del estudiantado en las clases (Hortigüela-Alcalá et al., 2021; James, 2018; Killian y Woods, 2021), su aplicación como criterios casi exclusivos puede generar consecuencias no deseadas. Tal como advierten Chernysh y Crossman (1994) y Krijgsman et al. (2017), la sobrevaloración del esfuerzo físico puede desmotivar a quienes poseen menores habilidades motrices, reproduciendo lógicas de exclusión. Por otro lado, Baghurst (2013) cuestiona que el énfasis exclusivo en actitudes pueda sustituir completamente la valoración del aprendizaje, debilitando el sentido formativo de la evaluación. Los hallazgos de este estudio reafirman esta tensión: si bien el profesorado valora la participación y el esfuerzo como indicadores clave del desempeño, emerge con claridad la necesidad de avanzar hacia sistemas evaluativos más equilibrados, que integren dimensiones actitudinales, procedimentales y conceptuales de forma articulada, promoviendo prácticas más inclusivas y pedagógicamente consistentes (Modell y Gerdin, 2022).

A pesar de los avances hacia una evaluación más inclusiva, persisten todavía algunas prácticas evaluativas tradicionales. Esto concuerda con los hallazgos de López-Pastor et al. (2012), quienes mencionan que, aunque existe un cambio hacia una evaluación alternativa en EF, todavía hay una fuerte presencia de enfoques tradicionales que continúan centrados en el rendimiento físico, sin considerar otras dimensiones del aprendizaje (Foglia et al., 2014). De igual forma, Talero y Guarnizo (2022) sostienen que la falta de instrumentos adecuados y de sistematización limita la posibilidad de implementar una evaluación verdaderamente alternativa en EF. Además, Prieto Ayuso (2016) destaca que, a pesar de la existencia de paradigmas evaluativos alternativos, la colaboración docente para la implementación de estos enfoques sigue siendo insuficiente, lo que dificulta un cambio

efectivo hacia una evaluación más comprensiva como lo sería un estilo actitudinal (Pérez, 2013) o propuestas con énfasis al esfuerzo y la superación (Lituma y Aldas, 2023).

6.4. Conclusiones estudio 3

Esta investigación permitió comprender cómo las experiencias evaluativas vividas durante la etapa escolar, junto con la formación inicial y la práctica docente, han influido en las prácticas actuales del profesorado de EF en Chile. El tránsito desde un enfoque centrado en el rendimiento físico hacia modelos que valoran el esfuerzo, la actitud y el proceso de aprendizaje se consolidó especialmente en la formación universitaria y el ejercicio profesional.

Como limitación, se reconoce que no todo el profesorado entrevistado facilitó sus instrumentos de evaluación lo que, si bien fue respetado éticamente, restringió una triangulación más profunda entre discurso y práctica.

Los hallazgos evidencian una incorporación progresiva de dimensiones actitudinales y formativas, impulsada por procesos reflexivos y por la necesidad de atender a la diversidad del estudiantado. Sin embargo, persisten tensiones, como la influencia de modelos tradicionales vividos en la etapa escolar y la exigencia de calificar numéricamente. Por lo tanto, se hace perentorio desarrollar nuevas investigaciones que, en el contexto chileno, pero también más allá, no solo profundicen en el conocimiento que tenemos de las conceptualizaciones y prácticas evaluativas del profesorado, sino que desarrollen intervenciones que cambien la realidad.



VNIVERSITAT E VALÈNCIA

CAPÍTULO 7

ESTUDIO 4

Estudio 4. La nota máxima está de moda. Análisis de las prácticas evaluativas en la asignatura de Educación Física y Salud chilena

Estado: Artículo aceptado en la revista Scientific Journal of School Sport, Physical Education and Psychomotricity SPORTIS.

Resumen

Este estudio analiza las prácticas evaluativas y calificativas del profesorado de Educación Física y Salud (EFyS) en Chile, mediante un diseño mixto secuencial exploratorio que integra una fase cualitativa basada en entrevistas semiestructuradas a docentes y un análisis cuantitativo de calificaciones oficiales correspondientes al período 2018–2022. El diseño metodológico fue mixto secuencial exploratorio. Los resultados cualitativos revelan que el profesorado tiende a concebir la evaluación como un proceso formativo que prioriza la participación, el esfuerzo y la actitud por sobre el rendimiento físico y a calificar con notas altas. En base a estos hallazgos se llevó a cabo la fase cuantitativa, la cual mostró una tendencia sostenida hacia calificaciones de excelencia, con una moda de 7,0 y un promedio nacional de 6,6 en 2022. Este fenómeno plantea interrogantes respecto a la relación entre calificación y aprendizaje efectivo, y sugiere la existencia de una cultura evaluativa que prioriza la inclusión y el reconocimiento del esfuerzo. Los hallazgos abren nuevas perspectivas sobre el sentido pedagógico de la calificación en EFyS, cuestionando enfoques tradicionales centrados en la medición motriz y destacando la necesidad de una evaluación más equitativa, comprensiva y coherente con los principios formativos de la asignatura. asignatura.

7.1. Metodología

La presente investigación se enmarca en un diseño mixto de tipo secuencial exploratorio (Creswell y Plano, 2011), compuesto por una fase cualitativa inicial que indagó las percepciones y prácticas evaluativas del profesorado de EFyS en contextos escolares chilenos, seguida de una fase cuantitativa centrada en el análisis de las calificaciones oficiales de la asignatura a nivel nacional. Este enfoque permite integrar la riqueza contextual de los métodos cualitativos con la sistematización numérica del análisis cuantitativo (Pole, 2009),

lo que favorece una comprensión más completa de fenómenos complejos en educación (Chaves, 2018; Şahin y Ozturk, 2019). En particular, en el ámbito de la EF, ha demostrado ser adecuado para abordar la evaluación y calificación desde una mirada amplia y contextualizada (Bagur-Pons et al., 2021; Castañer et al., 2013).

7.1.1. Fase cualitativa

7.1.1.1. Participantes

Las personas participantes en la fase cualitativa fueron 3 profesoras y 7 profesores de EFyS que, en el momento de la entrevista, se encontraban ejerciendo en diferentes establecimientos educativos de la ciudad de Talca, en la región chilena del Maule. Sus principales características pueden observarse en la Tabla 26. Dentro de los criterios de inclusión se consideró, como principal requisito, llevar al menos 1 año ejerciendo como docente de la asignatura de EFyS en cualquier contexto escolar y aceptar mediante un consentimiento informado participar voluntariamente en la investigación.

Tabla 26. Caracterización del profesorado entrevistado.

Participante	Género	Rango edad (años)	Años de experiencia docente
1	Mujer	50 a 59	32
2	Mujer	40 a 49	24
3	Hombre	30 a 39	11
4	Hombre	30 a 39	10
5	Hombre	30 a 39	10
6	Hombre	30 a 39	8
7	Hombre	30 a 39	7
8	Hombre	30 a 39	4
9	Hombre	20 a 29	4
10	Mujer	20 a 29	2

Elaboración propia

7.1.1.2. Técnica de regida y procedimiento

La primera fase cualitativa consistió en el desarrollo de entrevistas en profundidad y semiestructuradas, cuyo guion fue construido por el equipo de investigación. Las preguntas formuladas durante las entrevistas fueron cuidadosamente diseñadas para abordar diferentes categorías previamente identificadas en la literatura: conceptualización de evaluación desde la práctica, prácticas evaluativas y calificativas y evaluación sumativa (Evans, 2018; Hernández, 2014; Troncoso y Daniele, 2003).

Las entrevistas fueron realizadas de forma presencial y personalizada con cada una de las personas participantes, tuvieron una duración aproximada de 45 minutos, siendo registradas con grabadora de voz, para su posterior transcripción y análisis de unidades hermenéuticas, las cuales se describen en el siguiente apartado. Es importante mencionar que se utilizaron pseudónimos en las citas textuales en el apartado de resultados.

7.1.1.3. Análisis de datos fase cualitativa

Para el análisis de datos cualitativos provenientes de las entrevistas realizadas, se adoptó la metodología del análisis categorial inductivo, una técnica ampliamente reconocida en la literatura académica (Hodge y Sharp, 2016; Pérez-Samaniego et al., 2011). Esta estrategia de análisis permite una exploración profunda de los datos, comenzando con la identificación de las unidades de significado. Este proceso se realizó con cada una de las categorías previamente diseñadas que tenía la entrevista (Rodríguez et al., 2005). Una vez identificadas las unidades de significado, se asociaron con códigos que permitieran su agrupación temática en categorías. El resultado de la matriz final se puede observar en la tabla 27.

Tabla 27. Matriz de categorías elaboradas desde las entrevistas.

Categorías	Subcategorías	Códigos	Descripción
Conceptualización	Formativa	Progreso	Se describe la evaluación como un proceso, que debe considerar un seguimiento en un periodo de tiempo, para comunicar los avances.
		Proceso	
		Comunicada	
	Individualizada	Formativa	Se enfatiza en adecuar y considerar las características individuales del estudiantado en la evaluación, accesible para todos los niveles de competencia motriz.
		Más que calificar	
		Individual	
		Personalizada	
Proceso de calificación	Foco	Intereses estudiante	La concepción de evaluación se focaliza principalmente en la dimensión actitudinal, considerando el esfuerzo y la participación del estudiantado. En menor medida se mencionan las dimensiones procedimentales y conceptuales.
		Contexto	
		Adecuada	
	Frecuencia	Accesible	Se menciona que la evaluación está presente en todo momento, pero la calificación se genera 2 a 3 veces por semestre. Es decir, de 4 a 6 calificaciones anuales.
		Cualitativo	
		Esfuerzo	
Proceso de calificación	Foco	Comportamiento	La concepción de evaluación se focaliza principalmente en la dimensión actitudinal, considerando el esfuerzo y la participación del estudiantado. En menor medida se mencionan las dimensiones procedimentales y conceptuales.
		Información	
		Transición	
	Frecuencia	Conceptual	Se menciona que la evaluación está presente en todo momento, pero la calificación se genera 2 a 3 veces por semestre. Es decir, de 4 a 6 calificaciones anuales.
		Procedimental	
		Actitudinal	
Proceso de calificación	Frecuencia	Una por Unidad	Se menciona que la evaluación está presente en todo momento, pero la calificación se genera 2 a 3 veces por semestre. Es decir, de 4 a 6 calificaciones anuales.
		2 a 3 al semestre	
		Siempre	

Instrumentos	Registro clase a clase	Se mencionaron los registros clase a clase por observación, escalas de apreciación, rúbrica y listas de cotejo. Siempre pasando por un proceso de validación interno por UTP.
	Escala de apreciación Lista de cotejo Rúbricas Validados por UTP	
Dimensiones	Actitudinal	Se menciona con mayor frecuencia componentes de la dimensión actitudinal como elemento clave y de gran ponderación en la calificación. Se mencionan además las dimensiones procedimentales y conceptuales en menor medida.
	Comportamiento Participación Conceptuales Esfuerzo Superación Procedimentales	
Usos	Diagnóstico Formativo Calificar	Los usos atribuidos a la evaluación van desde el diagnóstico, seguimiento clase a clase, evaluaciones de procesos y calificación.
	Seguimiento de actitud Observación clase a clase	

Elaboración propia

7.1.2. Fase cuantitativa

La segunda fase de esta investigación corresponde al enfoque cuantitativo, adoptando un diseño longitudinal de tendencia, con alcance descriptivo (Delgado y Llorca, 2004). Esta etapa se desarrolló posteriormente al análisis cualitativo, con el propósito de complementar y ampliar la comprensión del fenómeno observado, permitiendo contrastar las percepciones y prácticas del profesorado con datos objetivos a nivel nacional.

7.1.2.1. Registros analizados

Se analizaron las calificaciones oficiales en la asignatura de EFyS de estudiantes desde primero básico hasta cuarto medio (6 a 18 años), proporcionadas por el Ministerio de Educación de Chile a través de una solicitud por transparencia pública. El periodo considerado fue de 2018 a 2022, delimitación que responde a criterios tanto normativos como metodológicos. El año 2018 se corresponde al año de promulgación del Decreto Exento N.º 67, que establece nuevas directrices para la evaluación escolar en Chile, mientras que 2022 coincide con el año en que se realizaron las entrevistas al profesorado. Las bases de datos utilizadas incluyen información sobre los promedios anuales de calificaciones por curso, nivel escolar y subsector, desagregados por cada una de las 16 regiones del país.

La tabla 28 presenta el número total de registros y establecimientos considerados en el análisis de cada año escolar.

Tabla 28. Bases de datos de calificaciones utilizadas en este estudio.

Año escolar	Número total de registros	Número de establecimientos
2018	30612736	9160
2019	32227180	9068
2020	25196613	9040
2021	32410770	8964
2022	32793724	8903

Elaboración propia

7.1.2.2. Análisis de datos fase cuantitativa

El análisis estadístico se hizo a partir de las bases de datos facilitadas por el Centro de Estudios del MINEDUC, que contienen las calificaciones reportadas de todos los establecimientos educacionales. Se utilizó estadística descriptiva para calcular la media, la desviación estándar y la moda de las calificaciones, para los promedios a nivel nacional, por región y por nivel de enseñanza utilizando el programa estadístico SPSS versión 25. Además, se graficó mediante un histograma el comportamiento de los promedios nacionales del año 2022.

De la totalidad de registros mencionados anualmente en la tabla 2, se hicieron los siguientes filtros:

- Por subsector, dejando únicamente la asignatura de EFyS.
- Por Nivel escolar; excluyendo los niveles de enseñanza de la educación básica y media para adultos.

7.1.3. Consideraciones éticas

Esta investigación fue aprobada por el Comité de Ética de la Universidad Autónoma de Chile (código N°CEC 38-21). En la fase cualitativa, se respetaron los principios de voluntariedad, confidencialidad y resguardo de la identidad. La participación del profesorado fue voluntaria, mediante consentimiento informado, y las entrevistas se realizaron de forma presencial, con una duración aproximada de 45 minutos. Fueron grabadas, transcritas y analizadas, utilizando pseudónimos para preservar la identidad de las personas participantes.

En la fase cuantitativa, se utilizaron bases de datos públicas obtenidas mediante la Ley de Transparencia a través del Ministerio de Educación. Aunque no contenían datos personales, se resguardó la privacidad institucional evitando referencias a establecimientos específicos, asegurando así el cumplimiento de los principios éticos en el manejo de datos secundarios.

7.2. Resultados

Conceptualizaciones evaluativas construidas desde la práctica educativa

Al albur de los datos, podemos afirmar que el profesorado entrevistado es capaz de hacer una nítida distinción entre evaluación y calificación, siendo la primera un concepto más amplio, como explica una docente:

"La evaluación para mí es analizar distintos tipos de cosas, ya sea evaluación de contexto, de contenidos, de comportamiento de los estudiantes. En general, el concepto de evaluación se puede aplicar para todo, uno puede evaluar cualquier situación. Es distinto cuando hablamos de una evaluación calificada, ahí ya estamos hablando de notas de los estudiantes, pero si me pregunta solo de evaluación, la evaluación se puede aplicar a cualquier aspecto que uno pretenda analizar en el fondo..." – Isidora

En relación con la concepción de la evaluación que tienen los y las participantes, parece existir cierto consenso en describir la evaluación como un proceso dinámico y formativo que va más allá de asignar calificaciones numéricas, valora el progreso individual de cada estudiante, y adopta metodologías y criterios de evaluación que garanticen la equidad y el desarrollo integral del estudiantado:

"Bueno, la evaluación en general para nosotros es más que el resultado, es valorar el proceso" - Antonia

En este orden de cosas, merece la pena hacer notar como se entiende la evaluación no tanto como un momento, sino como un proceso que, a su vez, debe ser formativo, tomando la opinión del estudiantado:

"Bueno, la evaluación en general, para nosotros es más que valorar el resultado por una nota, es evaluar el proceso..." - Antonia

"yo considero que la evaluación es un proceso completo, un proceso formativo que no tiene que ver solamente con el resultado final, sino que es un proceso que requiere de un inicio, basado en el interés del estudiante..." - Dante

"Para mí la evaluación es un proceso, si así tiene que ser, debería ser individual a cada estudiante..." - Felix

Prácticas evaluativas y calificativas

Las conversaciones sobre las prácticas evaluativas mostraron como el profesorado participante expresaba haber cambiado su forma de evaluar, dando pie a nuevas propuestas que se alejaban de un simple control del rendimiento físico:

“Por lo menos yo ahora, que soy jefe de departamento, traté de consensuar con los colegas el no calificar por rendimiento físico, por los test de condición física. Entonces evaluar rendimiento en el diagnóstico, no para calificarlo, sino para que nos sirva a nosotros para tener referencias de qué mejoraron en el semestre y en el año...” -Alan

En algunos casos, se ha mencionado la pandemia provocada por la COVID-19 como punto de inflexión a la hora de evaluar y calificar al alumnado:

“En años anteriores, cuando la prueba SIMCE estaba de moda, también calificábamos en base a test de condición física, pero no nos dio muchos resultados, finalmente porque la nota era muy baja, y al final era poco consecuente con nuestro pensamiento respecto a lo que queríamos lograr, y se aplicó hasta la pandemia...”-Dante

De una forma u otra, los datos reflejan una tendencia hacia el desarrollo de propuestas de evaluación más atingentes al paradigma formativo, y alejadas de la concepción técnica y de rendimiento. Con todo, algunos comentarios apuntan a la coexistencia de las dos formas de entender la evaluación:

“en el departamento de Educación Física del colegio existen 2 tipos de profesores, los profesores que son más antiguos y los profesores que son nuevos. Los más nuevos

trabajamos en todo el tema de evaluar lo procedimental y sobre todo lo actitudinal por sobre el rendimiento físico, que los niños clase a clase vayan adquiriendo aprendizajes, que vayan mejorando, esa es nuestra idea, pero de repente se nos quedan los profesores que son más antiguos que siguen calificando por capacidad física...” -Alan

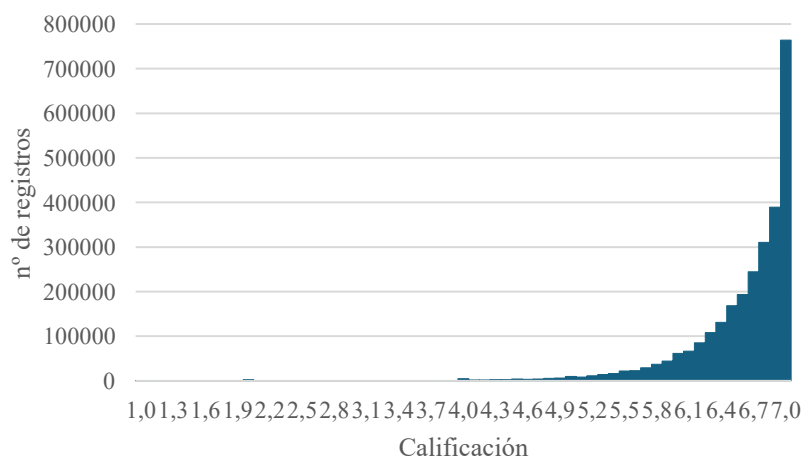
El profesorado entrevistado comenta de forma general que los promedios anuales de la asignatura se obtienen por el resultado de dos a tres calificaciones por semestre, es decir, una por unidad, en procedimientos validados por UTP, mediante instrumentos de evaluación (listas de cotejo, escalas de apreciación, rúbricas y registros de observación clase a clase). De forma unánime se comentan los rangos de promedios anuales en la asignatura, que van generalmente desde el 6,0 al 7,0, mencionando frases como, por ejemplo:

“Es común que se tengan promedios sobre 6,0, los que tienen bajas calificaciones es porque no participan” -Isidora

“Las evaluaciones se adaptan, para que todos puedan tener acceso a rendirlas de forma óptima...” -Dante

Las menciones sobre la calificación, expresadas por el profesorado participante en las entrevistas, explican de cierto modo las altas calificaciones observadas. Como se muestra en el histograma de la figura 1, aproximadamente el 28% de los escolares del país obtuvieron la calificación máxima en la asignatura de EF durante el año 2022.

Figura 6. Histograma calificaciones nacionales EFyS 2022. (Escala de 1 a 7).



Calificaciones de excelencia como tendencia nacional

En base a los comentarios de todo el profesorado entrevistado, que manifestaron rangos de calificaciones de excelencia, se contrastan dichas menciones con la tendencia nacional de los últimos 5 años.

En la tabla 29 se puede apreciar un comportamiento del promedio nacional en las calificaciones de la asignatura de excelencia, donde la tendencia de los cinco años de los promedios de todo el país, están por encima del 6,0, considerando que el 7,0 es la calificación máxima. La moda estadística en todos los años correspondió a la calificación máxima (7,0). La calificación media considerando los últimos 5 años, es un 6,4.

Tabla 29. Calificaciones de la asignatura de EFyS en función del nivel de enseñanza desde el 2018 al 2022.

	2018		2019		2020		2021		2022	
	N	M (DE)	N	M (DE)	N	M (DE)	N	M (DE)	N	M (DE)
Todos los niveles	2391419	6,44 (0,57)	2425677	6,48 (0,56)	1857483	6,18 (1,05)	2743759	6,31 (0,87)	2788981	6,60 (0,54)
Nivel enseñanza										
Básica	1947731	6,48 (0,52)	1977000	6,51 (0,51)	1399796	6,22 (1,02)	2008110	6,38 (0,81)	2028332	6,63 (0,47)
Media HC	350300	6,34 (0,71)	361522	6,38 (0,70)	366874	6,15 (1,09)	573955	6,20 (0,97)	596462	6,54 (0,64)
Media TP	93386	6,03 (0,83)	85630	6,12 (0,79)	90813	5,73 (1,20)	161683	5,85 (1,07)	164187	6,32 (0,78)

M= Media. DE=Desviación estándar; N=Población; Media HC= Enseñanza Media humanístico-científica; Media TP= Enseñanza Media Técnico profesional.

Considerando la extensión geográfica de Chile y su diversidad territorial, se estimó pertinente analizar la tendencia de las calificaciones por región. La tabla 30 muestra los promedios regionales en la asignatura de EFyS entre los años 2018 y 2022. En todos los casos, las calificaciones se mantienen consistentemente entre 6,0 y 6,7. De manera particular, se observa que el año 2022 concentra los promedios más altos en todas las regiones del país.

Tabla 30. Calificaciones de la asignatura de Educación Física y Salud en Chile según región geográfica.

Año	2018		2019		2020		2021		2022	
Región	N	M(DE)	N	M(DE)	N	M(DE)	N	M(DE)	N	M(DE)
Tarapacá	52622	6,44(0,59)	53979	6,47(0,58)	49690	6,11(1,10)	62569	6,28(0,88)	63054	6,62(0,52)
Antofagasta	90145	6,44(0,57)	92191	6,50(0,53)	85126	6,14(1,11)	103990	6,27(0,90)	106088	6,61(0,53)
Atacama	44405	6,42(0,65)	45901	6,47(0,61)	33174	6,26(1,04)	51364	6,31(0,89)	52989	6,58(0,59)
Coquimbo	113564	6,47(0,57)	116079	6,49(0,54)	93388	6,18(1,05)	133474	6,33(0,84)	137039	6,61(0,51)
Valparaíso	232856	6,45(0,58)	236890	6,48(0,56)	180197	6,18(1,07)	274139	6,35(0,85)	281009	6,60(0,54)
O'Higgins	130551	6,39(0,60)	132426	6,42(0,58)	90946	6,14(1,05)	151769	6,29(0,85)	155711	6,55(0,55)
Maule	145971	6,43(0,58)	147672	6,46(0,57)	103430	6,17(1,06)	171149	6,29(0,90)	175386	6,58(0,56)
Biobío	210047	6,51(0,53)	212413	6,54(0,52)	154256	6,35(0,95)	241252	6,41(0,81)	243748	6,64(0,50)
Araucanía	141228	6,46(0,54)	141620	6,47(0,54)	98571	6,22(0,98)	162381	6,33(0,83)	165239	6,59(0,53)
Los Lagos	121343	6,47(0,56)	121196	6,51(0,54)	54417	6,14(1,15)	138427	6,37(0,84)	139709	6,60(0,54)
Aysén	16529	6,46(0,55)	16354	6,46(0,57)	9644	6,36(1,02)	17881	6,34(0,88)	18178	6,56(0,60)
Magallanes	21829	6,58(0,50)	22345	6,59(0,48)	19621	6,46(0,88)	24445	6,46(0,76)	24836	6,69(0,47)
Metropolitana	917633	6,41(0,58)	930625	6,46(0,57)	771513	6,14(1,06)	1033580	6,27(0,90)	1045236	6,58(0,55)
Los Ríos	51645	6,48(0,54)	52460	6,53(0,53)	37970	6,20(1,04)	61804	6,33(0,85)	63419	6,64(0,49)
Arica	34682	6,55(0,54)	35642	6,59(0,52)	18283	6,13(1,07)	39862	6,33(0,87)	40043	6,68(0,50)
Ñuble	66367	6,51(0,53)	66450	6,54(0,52)	57257	6,29(0,96)	75662	6,44(0,76)	77297	6,66(0,48)

N=Población; M=Media; DE=Desviación estándar.

La importancia de la actitud y el esfuerzo en la calificación

La apertura hacia la consideración de dimensiones más allá de la puramente procedimental, se ve reflejada, por ejemplo, en la ponderación que le otorga el profesorado de EF a aspectos actitudinales en las calificaciones:

“...acá en el colegio, todas las calificaciones se ven afectadas por lo actitudinal y su participación clase a clase, se considera con el mismo peso, el 50% de ponderación en la nota de la Unidad...”-Alan

“... voy viendo cómo va mejorando, ahora, si el estudiante por ejemplo no intentó mejorar, yo veo a la hora de evaluación si está comprometido con la clase o quizás está preocupado de otra cosa. Ahí, obviamente, la evaluación no va a ser la misma que los demás, entonces ahí no va a obtener un 7...”-Isidora

En esta dirección, la dimensión conceptual de la enseñanza fue la menos señalada dentro del proceso de calificación. El énfasis entonces se encuentra tanto en las actitudes como en la dimensión procedimental:

“... por ejemplo, al momento de evaluar manipulación, los niños deben tener todas las guías pegadas y desarrolladas en su cuaderno. O sea, a los que tienen toda la parte

teórica, digamos completa, les adhiero un punto, dependiendo de del puntaje total de la evaluación...”- Isidora

“para la parte teórica, les hago una prueba al semestre a casi todos los cursos, pero todas las clases les hago preguntas para ver el manejo conceptual o lo que recuerdan de la clase pasada, esto no va calificado...” Jaime

La focalización de la evaluación en los aspectos procedimentales de la enseñanza, trae a colación la valoración del esfuerzo como elemento nuclear, no solo como aspecto a evaluar, sino también como hecho diferencial para la calificación del alumnado, como se desprende de la cita de Isidora. Ahondando en este hecho, el profesorado manifiesta en sus respuestas la importancia de dicho esfuerzo a la hora de calcular la calificación del alumnado:

“la calificación se consigue mucho más por el esfuerzo que ellos ponen, por la actitud que ellos colocan más que por la habilidad que ellos puedan tener, porque finalmente todo el currículum de nuestro colegio está planteado para que todos puedan lograrlo...”-Dante

“yo adapto esa evaluación hacia el niño que antes no saltaba, y ahora cuánto salta, si él presentó una mejoría en su condición o en su habilidad física, yo lo evalúo distinto, le puedo dar un puntaje extra, quizá por esfuerzo. O puedo ir ahí, tomando en cuenta en consideración algunos puntos que en las bases curriculares no aparecen”. -Isidora

Pero la reorientación de las prácticas evaluativas hacia elementos más ligados con las actitudes y el esfuerzo se asocia con dos aspectos importantes: por una parte, el hecho de animar al estudiantado a que participara activamente y mostrara respecto por sus compañeros y compañeras, la clase y el profesor, y, por la otra, a la necesidad de favorecer al estudiantado menos hábil motrizmente:

“yo soy bastante flexible en cuanto a la nota de los niños cuando veo progreso. Pero también soy muy estricta por la nota que tendrán por la conducta, que va incorporada dentro de los ítems. Se les conversa del impacto que tendrá en su calificación, y si no quieren participar o presentan mala actitud, la nota les baja bastante”. Isidora

“...en la misma clase les digo las cosas que voy a evaluar, donde siempre incorporo, por ejemplo, un ítem de convivencia y participación, y otro de conducta entonces eso también lo incorporo dentro de la evaluación...”-Felix

En cierta manera, la evaluación se utiliza como un incentivo a la participación del estudiantado, de forma que, es la motivación extrínseca la que entra en juego en forma de calificación:

“En la unidad de folklore, es un poco más complejo, porque no a todos les gusta bailar. Pero todos tienen que bailar. Y nosotros entendemos eso, también entendemos que no a todos les gusta bailar, entonces evaluamos el proceso, en el cual el estudiante se va poniendo notas, clase a clase por trabajar por participar en el baile, por seguir las instrucciones del profesor(a), por ensayar con su pareja o grupo. Se van poniendo una nota, entonces al estudiante se les explica que todos parten con un 7.0 la unidad, solo lo pueden bajar al no participar...”- Dante

“Los que no tienen notas altas, tiene más relación a conductas que no sean propias de la clase, o que no tenga participación. Pero tenemos estudiantes que se niegan a participar o que están lesionados, etc. nosotros también tenemos trabajo escritos para ellos, para que participen en la asignatura, aunque sea en la biblioteca...”-Felix

7.3. *Discusión*

La relación entre las políticas nacionales en educación, el currículum oficial y las prácticas evaluativas ha sido motivo de investigación en los últimos lustros (Svennberg, 2017). En el contexto chileno, el Decreto 67 (Ministerio de Educación, 2018) establece que la calificación final debe representar el grado de logro de los objetivos de aprendizaje, convirtiendo la nota en un símbolo del cumplimiento de propósitos educativos. Sin embargo, como advierte Zhu (2015), la relación entre calificación y aprendizaje real es problemática. En esta misma línea crítica, Santos-Guerra (1995) señala que suele privilegiarse la medición de la cantidad de aprendizajes por sobre la valoración de competencias superiores como el análisis, la crítica o la síntesis.

Desde la fase cualitativa de esta investigación, los relatos del profesorado entrevistado permiten identificar una concepción de la evaluación más procesual y abierta. Para ellos, el proceso de evaluación no se limita a la medición del desempeño físico observable, sino que incorpora variables como el esfuerzo, la participación y la actitud del estudiantado. Esta visión más comprensiva de la evaluación coincide con enfoques defendidos en la literatura reciente (Castillo-Retamal et al., 2024; Hortigüela et al., 2019; Moreno, 2016; Talanquer, 2015), que apuestan por valorar el progreso individual y el compromiso en lugar de centrarse exclusivamente en resultados cuantificables.

En este sentido, el profesorado entrevistado describió la utilización de instrumentos como rúbricas, listas de cotejo, escalas de apreciación y pautas de observación para registrar el progreso estudiantil, buscando transparentar el proceso evaluativo y otorgar objetividad a la calificación final. Estas herramientas deben ser validadas internamente por la UTP de cada establecimiento, reforzando la legitimidad del proceso, en línea con hallazgos de Portes-Junior et al. (2022), Castillo-Retamal et al. (2022) y Toro-Suaza y Chaverra-Fernández (2022). Esta tendencia no es exclusiva del contexto chileno; en Estados Unidos, por ejemplo, la NASPE ha promovido el uso de rúbricas estandarizadas para la evaluación motriz y cognitiva (Fox, 2012), alertando contra la evaluación "al ojo" que podría poner en riesgo la profesionalización docente (Matanin y Tannehill, 1994).

Dentro de esta mirada más integral de la evaluación, el profesorado considera que las bajas calificaciones no necesariamente indican un bajo aprendizaje, sino más bien reflejan una baja participación o compromiso en clase. Esta concepción también ha sido defendida por investigaciones que promueven evaluar dimensiones actitudinales junto a las competencias motrices (López-Pastor, 2006; Atienza et al., 2018; Cañadas, 2020; Toro-Suaza y Chaverra-Fernández, 2022), subrayando la necesidad de una evaluación auténtica, que priorice el aprendizaje consciente por sobre el rendimiento observable (Barba-Martín et al., 2020).

Sin embargo, el análisis debe situarse también en un marco estructural más amplio. La EF continúa enfrentando tensiones vinculadas a su estatus curricular en comparación con las asignaturas consideradas "core subjects", como Matemáticas o Lenguaje. Estudios recientes evidencian que esta situación impulsa a la asignatura a adoptar estrategias de evaluación cada

vez más estandarizadas y objetivas, en un intento por ganar legitimidad académica y posicionarse a la par de otras disciplinas del currículum escolar (Cruickshank et al., 2021; Pérez, 2021). Esta presión hacia la estandarización se traduce en la implementación de pruebas de condición física, instrumentos de medición del rendimiento y sistemas de evaluación que priorizan la objetividad cuantificable sobre procesos más formativos y humanistas (Harte et al., 2023; Zakaria et al., 2024). No obstante, esta orientación técnica entra en conflicto con el carácter humanizador de la Educación Física, cuyo sentido formativo abarca el desarrollo de habilidades sociales, valores, participación activa y experiencias motrices significativas (Sarni y Corbo, 2020). Además, revisiones recientes de las prácticas evaluativas en EF en Estados Unidos han evidenciado una falta de consistencia en los enfoques de evaluación, donde predomina aún la medición de habilidades físicas sobre la promoción de aprendizajes críticos y reflexivos (Killian y Woods, 2021). Todo lo anterior refleja un debate aún vigente en el campo educativo respecto a cómo equilibrar la necesidad de legitimidad curricular con la preservación de los principios formativos y emancipadores propios de la Educación Física escolar.

Pasando a los resultados de la fase cuantitativa, los análisis de las bases de datos oficiales (tablas 29 y 30) revelan una alta concentración de promedios entre 6,0 y 7,0 a nivel nacional y regional en la asignatura de EFyS. Esta distribución confirma lo anticipado en las entrevistas: la mayoría del estudiantado alcanza notas de excelencia, situándose abrumadoramente en los tramos superiores. Desde la crítica teórica, Biggs (2008) advierte que la calificación sumativa tiende a simplificar procesos de aprendizaje complejos, reduciéndolos a un número. De igual modo, Zhu (2015) y Santos Guerra (1993) cuestionan que la nota final refleje de forma fiable lo aprendido, destacando la distancia entre calificación y desarrollo real de competencias.

Estos datos también reflejan una cultura evaluativa que, si bien busca objetividad a través de instrumentos estandarizados, continúa evaluando en base a lógicas tradicionalistas, como lo alertan Fox (2012) y Matanin y Tannehill (1994), especialmente en Educación Física. Al integrar los resultados cualitativos y cuantitativos, puede sostenerse que en la asignatura de EFyS en Chile predomina un enfoque evaluativo que privilegia el esfuerzo, la participación y el compromiso, más que la medición estricta de la competencia física. Este

hallazgo se alinea con tendencias internacionales, donde se evidencia una progresiva incorporación de criterios actitudinales en la evaluación, aunque persisten elementos de corte tradicional. Así lo corroboran investigaciones en contextos latinoamericanos, como la de Toro-Suaza y Chaverra-Fernández (2022) en Colombia, donde, pese a algunos cambios, aún se privilegian mediciones centradas en el desempeño físico. Por su parte, en España, aunque se valoran los diagnósticos formativos y el mejoramiento de procesos de aprendizaje, persiste una orientación hacia la medición de resultados (Zubillaga y Cañadas, 2021).

En cuanto al significado de la calificación para el estudiantado, estudios recientes muestran que las altas calificaciones motivan la participación, en tanto que las bajas notas, especialmente si se asocian exclusivamente al rendimiento físico, pueden desmotivar (Kleinman, 1997; Krijgsman et al., 2017; Mackley et al., 2022; Sevil-Serrano et al., 2020). Así, las calificaciones ejercen una fuerte influencia en la vivencia escolar de las clases de EF (Martos-García y García-Puchades, 2023; Muñoz-Sepúlveda et al., 2025b).

Finalmente, este estudio confirma que el esfuerzo y la participación siguen siendo elementos clave en la evaluación en EFyS, en consonancia con otros contextos investigados (Baghurst, 2014; Hortigüela-Alcalá et al., 2021; James, 2018). No obstante, persiste el debate respecto de cuánto peso debe asignarse a estos componentes frente al aprendizaje motriz efectivo (Baghurst, 2013), y cuál es el equilibrio deseable en una evaluación formativa que permita tanto promover la equidad como garantizar el logro de aprendizajes significativos.

7.4. Conclusiones estudio 4

Esta investigación permitió identificar, en primer lugar, que el profesorado en ejercicio de la asignatura de EFyS en Chile comprende la evaluación como un proceso formativo que va más allá de la mera medición de habilidades motrices o condición física. La dimensión actitudinal: participación, respeto, conducta y esfuerzo, adquiere un valor preponderante al momento de calificar, por sobre los desempeños físicos o conceptuales. Esta prioridad se traduce en prácticas evaluativas centradas en la observación directa, listas de cotejo, rúbricas y escalas de apreciación, instrumentos que, cuando son utilizados para generar calificaciones, deben ser previamente validados por las Unidades Técnico-Pedagógicas de los establecimientos educativos.

En segundo lugar, el análisis de las calificaciones oficiales permitió concluir que estas se sitúan predominantemente en los rangos de excelencia, observándose una moda de 7,0 (nota perfecta) de forma recurrente en los últimos cinco años, y alcanzando un promedio nacional de 6,6 en el año 2022. Este fenómeno plantea interrogantes sobre si tales resultados reflejan efectivamente el logro pleno de los objetivos curriculares o si responden a otras dinámicas internas de la asignatura.

De este modo, se observa un cambio en el paradigma evaluativo del profesorado de EFyS chileno, alejándose progresivamente de un enfoque técnico-tradicional orientado al rendimiento, hacia una postura que prioriza la valoración del compromiso, la participación activa y el esfuerzo del estudiantado.

Sin embargo, una limitación de este estudio radica en que las entrevistas se realizaron en un contexto geográfico específico, lo que podría restringir la generalización de los hallazgos cualitativos a nivel nacional. Asimismo, el análisis de datos cuantitativos a partir de bases de datos agregadas no permitió vincular de manera directa las prácticas evaluativas de los docentes entrevistados con las calificaciones de su propio alumnado.

En perspectiva futura, sería relevante contrastar estos resultados con otras realidades internacionales, a fin de identificar si la alta concentración de notas de excelencia constituye un fenómeno particular del contexto chileno o si refleja una tendencia más generalizada en la evaluación de la Educación Física escolar a nivel latinoamericano o global.



VNIVERSITAT E VALÈNCIA

CAPÍTULO 8
CONCLUSIONES GENERALES

8.1. Conclusión general de la tesis

La presente tesis doctoral ha permitido construir una comprensión integral, situada y crítica del proceso evaluativo en la asignatura de EFyS en el sistema escolar chileno, mediante el desarrollo de cuatro estudios articulados en torno a las voces del estudiantado y del profesorado. A partir de esta estrategia metodológica, se abordó el fenómeno evaluativo en su complejidad, reconociendo su carácter contextual, su anclaje cultural y su dimensión ética y pedagógica.

Desde las experiencias del estudiantado (Estudios 1 y 2), emergen relatos que dan cuenta de prácticas evaluativas centradas principalmente en el rendimiento físico y la medición comparativa, muchas veces vividas como instancias de exposición, exclusión o juicio. Estas vivencias, atravesadas por la falta de criterios explícitos y por la ausencia de retroalimentación significativa, han moldeado una percepción ambivalente hacia la asignatura: por una parte, se valora su potencial motivador cuando el profesorado logra conectar con las necesidades e intereses del estudiantado; por otra, se problematiza su dimensión evaluativa cuando esta reproduce lógicas de competencia o estandarización. Además de incidir en el aprendizaje, estas prácticas han dejado huellas emocionales que afectan la relación con la asignatura e incluso con la actividad física en general. La evaluación no solo es interpretada como instrumento pedagógico, sino también como un dispositivo que puede reforzar desigualdades, jerarquizar cuerpos y limitar oportunidades de participación. La persistencia de estas prácticas, muchas veces heredadas o naturalizadas, revela la presencia de un currículum oculto que tensiona el sentido formativo de la evaluación escolar.

Por su parte, los estudios centrados en el profesorado (Estudios 3 y 4) evidencian un tránsito parcial desde modelos evaluativos tradicionales hacia prácticas más sensibles a la diversidad y al contexto del estudiantado. Las y los docentes entrevistados reconocen haber sido formados bajo paradigmas centrados en el rendimiento físico y el control corporal, pero manifiestan una progresiva resignificación de su rol evaluador, otorgando mayor importancia a dimensiones actitudinales como el esfuerzo, la participación y la disposición al aprendizaje. No obstante, este cambio no siempre se traduce en procedimientos sistemáticos, claros o

fundamentados, lo que evidencia una tensión entre las intenciones pedagógicas inclusivas y las herramientas concretas utilizadas para evaluar.

Este tránsito aparece atravesado por diferentes momentos del desarrollo profesional docente. En algunos casos, las transformaciones emergen en la propia formación inicial, mientras que, en otros, las y los profesores reconocen que sus prácticas evaluativas se modificaron tras observar directamente los efectos de la evaluación tradicional en el contexto escolar. Ello plantea un desafío central para las instituciones formadoras: instalar espacios de reflexión crítica sobre la evaluación desde etapas tempranas, tanto en asignaturas teóricas como en instancias de simulación o prácticas pedagógicas iniciales. Esta necesidad se acentúa al considerar que, en ciertos relatos, persiste la presencia de modelos universitarios centrados en el rendimiento físico o deportivo, lo que puede limitar las posibilidades de transformación real en el aula escolar.

El análisis cuantitativo complementario del Estudio 4 refuerza esta lectura al mostrar una alta concentración de calificaciones en los tramos más altos del sistema, siendo la nota máxima la moda anual entre 2018 y 2022. Si bien esta tendencia podría interpretarse como una valoración positiva del compromiso y la actitud del estudiantado, también plantea interrogantes respecto a la correspondencia entre las calificaciones otorgadas y los aprendizajes efectivamente logrados. En este sentido, el hallazgo sugiere un posible desbalance entre la función calificadora y la función formativa de la evaluación, así como la necesidad de avanzar hacia una cultura evaluativa que, sin renunciar al reconocimiento de lo actitudinal, incorpore criterios más claros, diversificados y coherentes con los propósitos pedagógicos de la asignatura.

En conjunto, los cuatro estudios que componen esta tesis doctoral permiten afirmar que la evaluación en la asignatura de EFyS en Chile se encuentra en un punto de inflexión. Si bien persisten prácticas ancladas en modelos tradicionales centrados en el rendimiento físico, la medición objetiva y la calificación como fin último, también emergen experiencias transformadoras que resignifican la evaluación como un proceso pedagógico situado, dialógico y formativo. Estas transformaciones no son lineales ni homogéneas; más bien,

revelan un escenario tensionado donde conviven la reproducción de esquemas heredados con búsquedas genuinas por una evaluación más justa, inclusiva y significativa.

Desde las voces del estudiantado, recogidas en los primeros dos estudios, se observa cómo la evaluación puede operar como un dispositivo de exclusión o como una oportunidad de aprendizaje, dependiendo del enfoque adoptado por el profesorado. Las experiencias narradas dan cuenta de prácticas evaluativas vividas con ansiedad, vergüenza o injusticia, especialmente cuando se aplican de manera estandarizada y sin contextualización. Por el contrario, se valoran aquellas instancias en que se incorpora la retroalimentación, la participación activa y criterios claros, capaces de generar sentido y motivación hacia la asignatura.

Los estudios centrados en el profesorado evidencian, por su parte, una toma de conciencia progresiva respecto al potencial pedagógico de la evaluación. Si bien se constata una intención por alejarse de la lógica exclusivamente técnica, aún existen vacíos en la formación inicial, dificultades para operacionalizar enfoques alternativos y escasa sistematización de criterios e instrumentos. La alta concentración de calificaciones máximas observada en el análisis cuantitativo del Estudio 4 ilustra esta tensión entre una evaluación que busca valorar lo actitudinal y una cultura pedagógica que todavía requiere mayor claridad, coherencia y fundamentación en sus prácticas.

Finalmente, esta tesis no solo aporta evidencia empírica sobre el estado actual de las prácticas evaluativas, sino que invita a repensar la evaluación como un acto profundamente pedagógico, que configura la experiencia escolar, la relación con el cuerpo y la manera en que se construyen sentidos sobre el aprender y el enseñar en EF. Reforzando la idea planteada por Santos Guerra (1995): “Dime cómo evalúas y te diré qué tipo de profesional y de persona eres”, la evaluación se proyecta aquí no solo como un procedimiento técnico, sino como una práctica ética, reflexiva y política, que requiere ser continuamente revisada en función de los contextos, las trayectorias y las subjetividades de quienes habitan la escuela.

8.2. Limitaciones y prospectiva

Esta investigación, al centrarse en un fenómeno tan complejo como la evaluación educativa, presenta ciertas limitaciones que es importante reconocer. Una de las principales

tiene que ver con el carácter cambiante de la educación, influido por factores culturales, institucionales y personales. La evaluación, en particular, se manifiesta de forma diversa en cada contexto, lo que dificulta establecer generalizaciones amplias a partir de los resultados obtenidos. Sin embargo, esta misma diversidad se puede considerar una fortaleza, ya que obliga a analizar las prácticas desde una mirada contextualizada, respetando las particularidades de cada realidad educativa.

Otro aspecto a considerar es que los datos obtenidos, tanto de estudiantes como de profesores, están profundamente marcados por sus experiencias personales, creencias y trayectorias. Esta subjetividad no debilita los hallazgos, sino que los enriquece, ya que permite comprender cómo se vive la evaluación desde quienes la protagonizan. No obstante, estos testimonios deben interpretarse con cautela, entendiendo que no buscan representar a todos los actores posibles, sino ofrecer una mirada situada y profunda sobre la temática.

Una limitación concreta de este estudio fue la diferencia temporal entre las dos etapas de recolección de datos. Las experiencias del estudiantado abarcan prácticas escolares vividas entre 2007 y 2019, mientras que las entrevistas al profesorado se realizaron entre 2021 y 2022, en un contexto postpandémico. Aunque la pandemia no fue parte del foco central del estudio, sí constituye un hecho relevante que pudo haber modificado las formas de evaluar en la práctica reciente.

Además, en el estudio centrado en el profesorado, no todos los docentes entrevistados compartieron sus instrumentos de evaluación. Esta situación limitó la posibilidad de hacer un análisis documental más completo y, por tanto, de contrastar con mayor profundidad los discursos sobre la evaluación con sus expresiones concretas. A ello se suma que en esa etapa no se realizaron observaciones de clases, lo cual podría haber aportado información valiosa para complementar los testimonios.

Pese a sus limitaciones, esta tesis doctoral abre rutas significativas para profundizar en la comprensión crítica del proceso evaluativo en la asignatura de EFyS. En primer lugar, se proyecta la necesidad de continuar investigando de manera sistemática y situada los sentidos que estudiantes y docentes atribuyen a la evaluación, reconociendo que esta práctica no solo cumple una función técnica, sino que se encuentra mediada por factores culturales, históricos,

políticos y subjetivos. Integrar nuevas técnicas de recolección de información, como el análisis audiovisual, la observación participante o estudios de caso prolongados, y diversificar los contextos escolares y universitarios permitiría enriquecer aún más la comprensión del fenómeno evaluativo, visibilizando matices que suelen quedar opacados por enfoques estandarizados o cuantitativos.

En segundo lugar, esta investigación invita a repensar la formación inicial y continua del profesorado como un espacio estratégico para transformar las prácticas evaluativas. Los hallazgos revelan una persistencia de modelos tradicionales y reproductivos, lo que plantea interrogantes sobre la manera en que las universidades están abordando la evaluación dentro de sus programas de pedagogía. Es fundamental revisar no solo los contenidos y metodologías de enseñanza de la evaluación, sino también las bases epistemológicas que las sustentan, interrogando las concepciones de aprendizaje, enseñanza y justicia educativa que las orientan. Asimismo, se vuelve imprescindible fomentar experiencias auténticas de evaluación formativa, autoevaluación, coevaluación y retroalimentación dialógica durante el proceso de formación, de modo que las y los futuros docentes se apropien críticamente de herramientas que les permitan valorar el proceso por sobre el producto.

En este sentido, se sugiere que las instituciones formadoras promuevan una actualización permanente respecto de la normativa vigente, como el Decreto N.º 67 (2018), los Estándares Pedagógicos y Disciplinarios, el MBE y la Política Nacional de Evaluación, pero que dicha actualización no se reduzca a una lectura instrumental, sino que sea abordada desde una perspectiva crítica y pedagógica, analizando su coherencia con los fines educativos y con las condiciones reales de las escuelas.

Una tercera línea de proyección se relaciona con el avance de las tecnologías digitales en el ámbito educativo. En particular, el auge de la inteligencia artificial generativa (IAg) plantea desafíos y oportunidades inéditos para el diseño y la implementación de procesos evaluativos. Esta tesis ofrece un punto de partida para futuras investigaciones que analicen el impacto de estas herramientas en la evaluación formativa, la retroalimentación automatizada, la personalización del aprendizaje y la agencia docente. Tales estudios

deberían considerar no solo la dimensión técnica de estas innovaciones, sino también sus implicancias éticas, políticas y pedagógicas.

Finalmente, se propone avanzar hacia investigaciones longitudinales, comparativas e interinstitucionales que permitan observar cómo evolucionan las culturas evaluativas en distintos tipos de escuelas, regiones, niveles educativos o trayectorias institucionales. Asimismo, sería relevante explorar cómo las políticas públicas, los lineamientos institucionales y las culturas organizacionales inciden en las decisiones evaluativas que toman las y los docentes en sus prácticas cotidianas. Un punto crítico que emerge de esta investigación es la disociación entre lo que se evalúa y lo que finalmente se califica, observándose una tendencia generalizada a otorgar calificaciones en rangos de excelencia que tensiona la validez de los procedimientos utilizados y cuestiona el sentido de la calificación en EFyS. Este fenómeno requiere ser indagado en profundidad, especialmente en contextos donde se promueve una evaluación orientada al proceso, la mejora y la justicia educativa. En este escenario, cobra especial relevancia el proceso de actualización de las Bases Curriculares que está en curso en el sistema escolar chileno durante el año 2025, el cual incluye la asignatura de EF. Este hito representa una oportunidad para generar nuevas investigaciones centradas en la interpretación, implementación y evaluación del nuevo currículum, así como en los desafíos que este plantea para la formación docente y la construcción de una cultura evaluativa coherente con los principios formativos. En suma, las proyecciones de esta tesis se orientan hacia el fortalecimiento de una evaluación comprometida con el aprendizaje, la equidad y el desarrollo integral del estudiantado.



VNIVERSITAT DE VALÈNCIA

REFERENCIAS

Referencias

- Aarskog, E. (2020). No assessment, no learning: Exploring student participation in assessment in Norwegian physical education (PE). *Sport, Education and Society*, 26(8), 875–888. <https://doi.org/10.1080/13573322.2020.1791064>
- Aartun, I., Walseth, K., Standal, Ø. F., & Kirk, D. (2022). Pedagogies of embodiment in physical education - A literature review. *Sport, Education and Society*, 27(1), 1–13. <https://doi.org/10.1080/13573322.2020.1821182>
- Abós, A., Burgueño, R., García-González, L., & Sevil-Serrano, J. (2021). Influence of internal and external controlling teaching behaviors on students' motivational outcomes in physical education: Is there a gender difference? *Journal of Teaching in Physical Education*, 41(3), 502–512. <https://doi.org/10.1123/jtpe.2020-0316>
- Adolfo, M. G. V., Belocura, C. V., Beltran, E. V., & Derasin, L. M. C. (2023). Insights of high school students learning folk dance during physical education classes. *International Journal of Science and Management Studies*, 6(3), 107–113. <https://doi.org/10.51386/25815946/ijsms-v6i3p107>
- Agencia de Calidad de la Educación. (2018). *Nuevo Sistema Nacional de Evaluación de Aprendizajes: La evaluación al servicio de los aprendizajes*. Santiago, Chile. <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/17467>
- Aguilar-Farias, N., Cortinez-O’Ryan, A., Sadarangani, K. P., Oetinger, A. von, Leppe, J., Valladares, M., Balboa-Castillo, T., Cobos, C., Lemus, N., Walbaum, M., & Cristi-Montero, C. (2016). Results from Chile’s 2016 report card on physical activity for children and youth. *Journal of Physical Activity and Health*, 13(s2), S117–S123. <https://doi.org/10.1123/jpah.2016-0314>
- Ahumada, P. A. (2005). La evaluación auténtica: Un sistema para la obtención de evidencias y vivencias de los aprendizajes. *Revista Perspectiva Educacional*, 45, 11–24. <https://www.redalyc.org/articulo.oa?id=333329100002>

- Aiello, M., & Fernández-Carneiro, T. (2021). Desafíos de la gestión para la formación de investigadores en educación: Planteamientos desde la perspectiva de género en educación física. *Lúdica Pedagógica*, 1(33), 29–35. <https://doi.org/10.17227/ludica.num33-13209>
- Alfrey, L., & Gard, M. (2014). A crack where the light gets in: A study of health and physical education teachers' perspectives on fitness testing as a context for learning about health. *Asia-Pacific Journal of Health, Sport and Physical Education*, 5(1), 3–18. <https://doi.org/10.1080/18377122.2014.867790>
- Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice*, 20(1), 20–34. <https://doi.org/10.1080/0969594X.2012.736364>
- Almerich, G., Díaz-García, I., Cebrián-Cifuentes, S., & Suárez-Rodríguez, J. (2018). Dimensional structure of 21st century competences in university students of education. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 24(1), 1–18. <https://doi.org/10.7203/RELIEVE.24.1.12548>
- Álvarez Cisternas, M., Torres Orellana, B., & Medina Guajardo, I. (2022). Retroalimentación efectiva con estímulos positivos desde lo emocional. *EDU Review. International Education and Learning Review*, 10, 1–9. <https://doi.org/10.37467/gkarevedu.v10.3171>
- Álvarez Rivera, L., Cuevas Ferrera, R., Lara Pot, A., & González Hernández, J. (2015). Diferencias del autoconcepto físico en practicantes y no practicantes de actividad física en estudiantes universitarios. *Cuadernos de Psicología Del Deporte*, 15(2), 27–34. <https://doi.org/10.4321/s1578-84232015000200004>
- Alvariñas-Villaverde, M., & Pazos-González, M. (2018). Estereotipos de género en Educación Física, una revisión centrada en el alumnado. *Revista Electrónica de Investigación Educativa*, 20(4), 154-163. <https://doi.org/10.24320/redie.2018.20.4.1840>
- Andrade, H., & Heritage, M. (2017). *Using formative assessment to enhance learning, achievement, and academic self-regulation* (1st ed.). Routledge. <https://doi.org/10.4324/9781315623856>

- Angelini Doffo, M. (2016). Estudio sobre la evaluación formativa y compartida en la formación docente en inglés. *Actualidades Investigativas en Educación*, 16. <https://doi.org/10.15517/aie.v16i1.22614>
- Anijovich, R., Comilloni, A., Cappelletti, G., Hoffman, J., Katzkowicz, R., & Mottier, L. (2010). *Evaluación significativa* (1ª ed.). Paidós.
- Anijovich, R. (2017). La evaluación formativa en la enseñanza superior. *Voces de la educación*. 2 (1) pp. 31-38. <https://hal.science/hal-02540341>
- Añazco-Martínez, L. A., Gallardo-Fuentes, F. J., & Carter-Thuillier, B. I. (2024). Evaluación formativa y compartida en educación física escolar: una revisión sistemática de las experiencias/publicaciones/estudios recientes en América Latina. *Revista Internacional de Investigación en Ciencias Sociales*, 20(1), 51-72. <https://doi.org/10.18004/riics.2024.junio.51>
- Aranda, M. T. F., & Llontop, L. R. A. (2023). Política de evaluación y su implicancia en la labor docente de una comunidad educativa peruana. *Ensaio: Avaliação e Políticas Públicas em Educação*, 31(121), e0234017. <https://doi.org/10.1590/s0104-40362023003104017>
- Argüello, A. (2014). La perspectiva biográfica en la investigación educativa: Una mirada a sus antecedentes, tendencias y posibilidades. *Estudios Pedagógicos*, 40(1), 293–308. <http://dx.doi.org/10.4067/S0718-07052014000100018>
- Asenjo-Paredes, C., Gallardo-Fuentes, F., Carter-Thuillier, B., & Peña-Troncoso, S. (2024). Concepciones y prácticas evaluativas en la formación inicial del profesorado en Educación Física. *Revista Ciencias de la Actividad Física UCM*, 25(1), 1–17. <https://doi.org/10.29035/rcaf.25.1.9>
- Asociación Médica Mundial. (2024). Declaración de Helsinki de la AMM – *Principios éticos para las investigaciones médicas en seres humanos*. <https://www.wma.net/es/policies-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/>

- Asún-Dieste, S., & Guáu Carrera, M. (2023). Use of formative assessment in Physical Education teacher education in Secondary School: a case study. *Cultura Ciencia y Deporte*, 18(55), 171–190. <https://doi.org/10.12800/ccd.v18i55.1918>
- Atienza Gago, R. (2022). *Avaluació formativa i compartida en la formació inicial del professorat d'educació física. Una aproximació des de la perspectiva de l'alumnat* [Tesis doctoral, Universitat de València]. Repositorio RODERIC. <https://hdl.handle.net/10550/83401>
- Atienza, R., Valencia, A., & Devís, J. (2018). Experiencias de Evaluación en Educación Física. Una Aproximación desde la Formación Inicial del Profesorado. *Estudios Pedagógicos*, 44(2), 127–147. <https://doi.org/10.4067/S0718-07052018000200127>
- Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology: A cognitive view* (2nd ed.). New York: Holt, Rinehart and Winston.
- Baghurst, T. (2013). How physical education teachers undermine the profession by grading effort and participation. *Oklahoma AHPERD*, 50(3), 51–55. <https://api.semanticscholar.org/CorpusID:166925788>
- Baghurst, T., & Mwavita, M. (2014). Evaluation, rationale, and perceptions regarding fitness testing in physical education teacher education programs. *The Global Journal of Health and Physical Education Pedagogy*, 3(4), 349–365.
- Bagur-Pons, S., Rosselló-Ramon, M. R., Paz-Lourido, B., & Verger, S. (2021). El enfoque integrador de la metodología mixta en la investigación educativa. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 27(1), art. 3. <https://doi.org/10.30827/relieve.v27i1.21053>
- Barba-Martín, R. A., Hortigüela Alcalá, D., & Pérez-Pueyo, Á. (2020). Evaluar en Educación Física: análisis de la tensiones existentes y justificación del empleo de la evaluación formativa y compartida. *Educación Física y Deporte*, 39(1), 23–47. <https://doi.org/10.17533/udea.efyd.v39n1a03>
- Barrientos, E. J., López-Pastor, V. M., Lorente-Catalán, E., & Kirk, D. (2023). Challenges with using formative and authentic assessment in physical education teaching from experienced

- teachers' perspectives. *Curriculum Studies in Health and Physical Education*, 14(2), 109–126. <https://doi.org/10.1080/25742981.2022.2060118>
- Barrientos-Hernán, E., López-Pastor, V., & Pérez-Brunicardi, D. (2020). Evaluación Auténtica y Evaluación Orientada al Aprendizaje en Educación Superior. *Revista Iberoamericana de Evaluación Educativa*, 13(2). <https://doi.org/10.15366/rie2020.13.2.004>
- Basurto-Mendoza, S. T., Moreira-Cedeño, J. A., Velásquez-Espinales, A. N., & Rodríguez-Gámez, M. (2021). Autoevaluación, Coevaluación y Heteroevaluación como enfoque innovador en la práctica pedagógica y su efecto en el proceso de enseñanza-aprendizaje. *Polo del Conocimiento*, 6(3), 828–845. <https://polodelconocimiento.com/ojs/index.php/es/article/view/2408>
- Beltrán-Carrillo, V. J., Devís-Devís, J., & Brown, D. H. K. (2012). When physical activity participation promotes inactivity: Negative experiences of Spanish adolescents in physical education and sport. *Youth & Society*, 44(1), 3–23. <https://doi.org/10.1177/0044118X10388262>
- Beltrán-Carrillo, V. J., & Devís, J. D. (2019). El pensamiento del alumnado inactivo sobre sus experiencias negativas en educación física: los discursos del rendimiento, salutismo y masculinidad hegemónica. *RICYDE. Revista Internacional de Ciencias del Deporte*, 15(55), 20-34. <https://doi.org/10.5232/ricyde2019.05502>
- Bendezú Monge, T. (2025). Evaluación formativa para mejorar los aprendizajes: Revisión sistemática. *Horizontes. Revista de Investigación en Ciencias de la Educación*, 9(36), 681–698. <https://doi.org/10.33996/revistahorizontes.v9i36.945>
- Beriche, M. & Medina, Z. (2021). La evaluación formativa: su implementación y principales desafíos en el contexto de la escuela y la educación superior. *Educación*, 27(2), 201–208. <https://doi.org/10.33539/educacion.2021.v27n2.2433>
- Bermeo-Yaffar, F., Hernández-Mosqueda, J., & Tobón-Tobón, S. (2016). Análisis documental de la V heurística mediante la cartografía conceptual. *Ra Ximhai*, 12(6), 103–121. <https://doi.org/10.35197/rx.12.01.e3.2016.05.fb>

- Bernard, E., & González-Moreno, P. A. (2020). Evaluación auténtica como alternativa para la mejora en el proceso de enseñanza-aprendizaje de educación física en escuelas primarias. *Revista Ciencias de la Actividad Física UCM*, 21(2), 1–19. <https://doi.org/10.29035/rcaf.21.2.3>
- Bezeau, D., Turcotte, S., Desbiens, J. F., Spallanzani, C., Roy, M., Vanderclayen, F., & Beaudoin, S. (2023). Physical education teachers' assessment practices in health education. *Physical Education and Sport Pedagogy*, 1–13. <https://doi.org/10.1080/17408989.2023.2230280>
- Biesta, G. (2009). Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, 21(1), 33–46. <https://doi.org/10.1007/s11092-008-9064-9>
- Bixio, C. (2006). *Evaluación de los aprendizajes en el nivel medio: Aportes para pensar las prácticas*. Buenos Aires: Paidós.
- Bizarro, W., Paucar, P. y Chambi-Mescoco, E. (2021). Evaluación formativa: una revisión sistemática de estudios en aula. Horizontes. *Revista de Investigación en Ciencias de la Educación*, 5(19), 872–891. <https://doi.org/10.33996/revistahorizontes.v5i19.244>
- Black, P. J., & Wiliam, D. (2003). 'In praise of educational research': Formative assessment. *British Educational Research Journal*, 29(5), 623–637. <https://doi.org/10.1080/0141192032000133721>
- Black, P. J., & Wiliam, D. (2018). Classroom Assessment and Pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25, 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. McGraw-Hill.
- Bolívar, A. (2002). ¿De nobis ipsis silemus?: Epistemología de la investigación biográfico-narrativa en educación. *REDIE: Revista Electrónica de Investigación Educativa*, 4(1), 1–26. <https://redie.uabc.mx/redie/article/view/49/91>

- Bores-García, D., Hortigüela-Alcalá, D., Hernando-Garijo, A., & González-Calvo, G. (2021). Análisis de la motivación del alumnado hacia la expresión corporal a través del uso de la evaluación formativa y compartida. *Retos: Nuevas tendencias en Educación Física, Deporte y Recreación*, 40, 198–208. <https://doi.org/10.47197/retos.v1i40.83025>
- Borghouts, L. B., Slingerland, M., & Haerens, L. (2017). Assessment quality and practices in secondary PE in the Netherlands. *Physical Education and Sport Pedagogy*, 22, 473–489. <https://doi.org/10.1080/17408989.2016.1241226>
- Bourdieu, P., & Passeron, J.-C. (2007). *La reproducción: Elementos para una teoría del sistema de enseñanza* (2.^a ed.). Ediciones Laia.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Broadfoot, P. (2007). *An introduction to assessment*. Continuum.
- Brown, S. (2015). Perspectivas internacionales sobre la práctica de la evaluación en Educación Superior. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 21(1), art. ME7. <https://doi.org/10.7203/relieve.21.1.6403>
- Bryant, A., & Charmaz, K. (Eds.). (2007). *The SAGE Handbook of Grounded Theory*. Sage Publications. Reino Unido.
- Buitrago, T., Cabezas, M., Castillo, J., Moyano, A., & Pinzón, M. (2018). Evaluación auténtica: Un camino hacia la transformación de las prácticas pedagógicas. *Revista de Educación y Desarrollo Social*, 12, 74–89. <https://doi.org/10.18359/reds.3359>
- Bulut, O., Beiting-Parrish, M., Casabianca, J. M., Slater, S. C., Jiao, H., Song, D., ... & Liu, J. X. (2024). The Rise of Artificial Intelligence in Educational Measurement: Opportunities and Ethical Challenges. *Journal of Educational Measurement and Evaluation*, 5(3), 2024, Article 3. <https://arxiv.org/abs/2406.18900>
- Calatayud Salom, M. A. (2019). Una oportunidad para avanzar hacia la evaluación auténtica en Educación Física. *Retos, Nuevas tendencias en Educación Física, deporte y recreación*, 36, 259–265. <https://doi.org/10.47197/retos.v36i36.67540>

- Calatayud Salom, M. (2021). Metáforas evaluativas del alumnado universitario de Educación Física. Necesidad de un cambio radical. *Retos, Nuevas tendencias en Educación Física, deporte y recreación*, 41, 854–865. <https://doi.org/10.47197/retos.v41i0.82706>
- Caldeiro, G. (2024). Inteligencia artificial generativa y educación: Hacia un nuevo paradigma. *El Faro. Revista Digital de Docencia Universitaria*, 1(1), 22–43. <https://revistaelfaro.uflo.edu.ar/index.php/elfaro/article/view/15>
- Calvo, G. G. (2013). *Evolución de la identidad profesional de un docente novel de Educación Física: análisis a partir de la reflexión de sus experiencias personales y de su propia práctica* (Doctoral dissertation, Universidad de Valladolid). <https://doi.org/10.35376/10324/2817>
- Camacho-Miñano, M. J., Fernández-García, E., & Rodríguez-Galiano, M. I. (2006). Imagen Corporal y práctica de actividad física en las chicas adolescentes: Incidencia de la modalidad deportiva. *RICYDE. Revista Internacional de Ciencias Del Deporte*, 2(3), 1–19. <https://doi.org/10.5232/ricyde2006.00301>
- Camacho-Miñano, M. J., & Girela-Rejón, M. J. (2017). Evaluación de una propuesta formativa sobre género en Educación Física para estudiantes de Ciencias de la Actividad Física y el Deporte. *Cultura, Ciencia y Deporte*, 12(36), 195–202. <https://doi.org/10.12800/ccd.v12i36.950>
- Campbell-Patton, C., & Patton, M. (2017). Book Review: The Future of Evaluation: Global Trends, New Challenges, Shared Perspectives. *American Journal of Evaluation*, 38(3), 453–458. <https://doi.org/10.1177/1098214017698036>
- Caniupán Alcayaga, K., & Valenzuela Gómez, A. (2025). Uso de resultados de brecha de género en evaluaciones educativas a gran escala (SIMCE y DIA) por parte de los establecimientos educacionales chilenos. *Revista Realidad Educativa*, 5(1), 1–18. <https://doi.org/10.38123/rre.v5i1.473>
- Cañadas, L. (2020). Evaluación formativa en el contexto universitario: Oportunidades y propuestas de actuación. *Revista Digital de Investigación en Docencia Universitaria*, 14(2), 9–22. <https://doi.org/10.19083/ridu.2020.1214>

- Cañadas, L., Valencia-Peris, A., & Sevil-Serrano, J. (2021). Capítulo 2.8. Cómo aplicar la evaluación formativa para favorecer la motivación y el aprendizaje en educación física. En I. Aguaded (Ed.), *Innovación en educación física y deporte: Nuevas tendencias y propuestas didácticas* (pp. 159–176). Editorial Universidad de Zaragoza. <https://doi.org/10.26754/uz.978-84-18321-22-1>
- Carbajosa, D. (2011). Debate desde paradigmas en la evaluación educativa. *Perfiles Educativos*, 33(132), 183–192. <https://doi.org/10.22201/iisue.24486167e.2011.132.24903>
- Cardon, P., Fleischmann, C., Aritz, J., Logemann, M., & Heidewald, J. (2023). The Challenges and Opportunities of AI-Assisted Writing: Developing AI Literacy for the AI Age. *Business and Professional Communication Quarterly*, 86(3), 257-295. <https://doi.org/10.1177/23294906231176517>
- Carless, D., & Douglas, K. (2011). Stories as personal coaching philosophy. *International Journal of Sports Science and Coaching*, 6(1), 1-12. <https://doi.org/10.1260/1747-9541.6.1.1>
- Castanedo Alonso, J., & Capllonch Bujosa, M. (2017). La Evaluación Formativa y Compartida en el Modelo Inclusivo Retos Individuales con Responsabilidad Compartida. *Revista Infancia, Educación y Aprendizaje*, 3. <https://doi.org/10.22370/ieya.2017.3.2.710>
- Castañer, M., Camerino, O. y Anguera, M. (2013). Mixed methods in the research of sciences of physical activity and sport. *Apunts. Educación Física y Deportes*, 112, 31–36. [http://dx.doi.org/10.5672/apunts.2014-0983.es.\(2013/2\).112.01](http://dx.doi.org/10.5672/apunts.2014-0983.es.(2013/2).112.01)
- Castillo, J. (2021). *El uso de los resultados de las pruebas de evaluación de los aprendizajes en el planeamiento de las políticas educativas en Chile: Informe nacional*. Buenos Aires: IIPE-UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000379596>
- Castillo, C. M., Payeras, P. S., Prat, S. S., Soler, A. V., Rivera, C. F., Knijnik, J., & Villouta, P. L. (2023). Creencias y prácticas sobre la perspectiva de género en el profesorado de Pedagogía en Educación Física en Chile. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, (47), 969-977. <https://doi.org/10.47197/retos.v47.96869>

- Castillo Retamal, F., Matus-Castillo, C., & Cornejo, M. (2021). La perspectiva de género en la formación inicial docente en la educación física chilena. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, 40, 830-847. <https://doi.org/10.47197/retos.v1i40.83082>
- Castillo-Retamal, F., Sánchez-Quiroz, A., Montepil-Núñez, R., Quintana-Moya, L., Rojas-Román, M., & Villanueva-Carrasco, E. (2024). Evaluación en Educación Física: percepción del profesorado desde el territorio escolar. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, 52, 610–622. <https://doi.org/10.47197/retos.v52.101527>
- Castillo-Retamal, F., Souza de Carvalho, R., Bássoli de Oliveira, A. A., Matias de Souza, V. de F., Barbosa Anversa, A. L., & Pereira da Silva Júnior, A. (2022). Evaluación en Educación Física escolar: discusiones a partir de la formación de profesores. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, 46, 179–189. <https://doi.org/10.47197/retos.v46.93736>
- Cerón, C., & Echeverría, E. (2020). La evaluación docente en Chile en los últimos 5 años. *Anuario Digital de Investigación Educativa*, 2, 426–438. <https://revistas.bibdigital.uccor.edu.ar/index.php/adiv/article/view/4241>
- Charmaz, K. (2014). *Constructing Grounded Theory* (2ª ed.). Sage Publications. Reino Unido.
- Chaverra Fernández, B. E., & Hernández Álvarez, J. L. (2019). La acción evaluativa en profesores de Educación Física: una investigación multi-casos. *Revista Iberoamericana de Evaluación Educativa*, 12(1), 199–215. <https://doi.org/10.15366/riee2019.12.1.012>
- Chaves Montero, A. (2018). *La utilización de una metodología mixta en investigación social*. En Kenneth Delgado; Santa Gadea; Walter Federico Gadea y Sara Vera-Quiñonez (Coords.), *Rompiendo barreras en la investigación* (1ª ed. en español, pp. 164–184). UTMACH
- Chernysh, K., & Crossman, J. (1994). The perceptions of students regarding grading in physical education. *The Canadian Association for Health, Physical Education & Recreation*, 60(1), 21–24. <https://www.cabidigitallibrary.org/doi/full/10.5555/19941804437>

- Chihuailaf-Vera, M. L., Ferro, E. F., Cid, F. M., & Jiménez, R. G. (2024). Estereotipos de género en la práctica de ejercicio físico y deporte en estudiantes universitarios de la carrera de Pedagogía en Educación Física en Chile. *Retos: nuevas tendencias en educación física, deporte y recreación*, 55, 1-10. <https://dialnet.unirioja.es/servlet/articulo?codigo=9480660>
- Chng, L. S., & Lund, J. (2018). Assessment for learning in physical education: The what, why and how. *Journal of Physical Education, Recreation and Dance*, 89, 29-34. <https://doi.org/10.1080/07303084.2018.1503119>
- Cisternas-Villalobos, C., Batista, P., Merellano-Navarro, E., & RibeiroSilva E. (2023). Identidad Profesional Docente en la Educación Física: revisión integrativa de literatura. *Ebalonmano*, 19(3), 276-290. <https://doi.org/10.17398/1885-7019.19.276>
- Claver, F., Martínez-Aranda, L. M., Conejero, M., & Gil-Arias, A. (2020). Motivation, discipline, and academic performance in physical education: A holistic approach from achievement goal and self-determination theories. *Frontiers in Psychology*, 11, 1808. <https://doi.org/10.3389/fpsyg.2020.01808>
- Collier, D. (2011). Increasing the Value of Physical Education. *Journal of Physical Education, Recreation & Dance*, 82(7), 38–41. <https://doi.org/10.1080/07303084.2011.10598656>
- Conde-Rojas, F. J., Martina Arroyo, M., & Aniceto Vargas, P. F. (2022). Análisis del concepto evaluación educativa mediante cartografía conceptual, camino hacia la calidad. *Ciencia Latina Revista Científica Multidisciplinar*, 6(1), 2867-2888. https://doi.org/10.37811/cl_rcm.v6i1.1694
- Connelly, F. M., & Clandinin, D. J. (1990). Stories of experience and narrative inquiry. *Educational Researcher*, 19(5), 2–14. <https://doi.org/10.3102/0013189X019005002>
- Connelly, F. M., & Clandinin, D. J. (1995). *Relatos de experiencia e investigación narrativa*. En J. Larrosa et al. (Eds.), *Déjame que te cuente. Ensayos sobre narrativa y educación* (pp. 11–59). Laertes.
- Corbin, C. B. (2010). Texas Youth Fitness Study: A Commentary. *Research Quarterly for Exercise and Sport*, 81(sup3), S75–S78. <https://doi.org/10.1080/02701367.2010.10599696>

- Corbin, J., & Strauss, A. (2014). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage Publications.
<https://www.scirp.org/reference/referencespapers?referenceid=3252683>
- Cordero Arroyo, G., Luna Serrano, E., & Patiño Alonso, N. (2013). La evaluación docente en educación básica en México: panorama y agenda pendiente. *Sinéctica Revista electrónica de Educación*, 41, 2–19. <https://www.scielo.org.mx/pdf/sine/n41/n41a8.pdf>
- Córdoba Peralta, A. L., & Lanuza Saavedra, E. M. (2021). Breve revisión conceptual sobre la evaluación de los resultados académicos en el sistema educativo. *Revista Científica Estelí*, 36–48. <https://doi.org/10.5377/farem.v0i0.11606>
- Cornejo, M., & Salas, N. (2011). Methodological rigor and quality: A challenge to qualitative social research. *Psicoperspectivas*, 10(2), 12–34. <https://doi.org/10.5027/psicoperspectivas-vol10-issue2-fulltext-144>
- Corrales-Perea, Á., & Espada, M. (2022). Motivación y percepción del alumnado en los estilos de enseñanza mando directo y resolución de problemas en educación física. *Revista Electrónica Educare*, 26(3), 17–34. <https://doi.org/10.15359/ree.26-3.2>
- Coterón López, J., Franco Álvarez Álvarez, E., Pérez Tejero, J., & Sampedro, J. (2013). Clima motivacional, competencia percibida, compromiso y ansiedad en Educación Física: Diferencias en función de la obligatoriedad de la enseñanza. *Revista de Psicología del Deporte*, 22(1), 151–157. <http://hdl.handle.net/11531/21761>
- Cresswell, M. J. (2003). *Heaps, prototypes and ethics: The consequences of using judgements of student performance to set examination standards in a time of change*. University of London Institute of Education.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Cruikshank, V., Hyndman, B., Patterson, K. y Kebble, P. (2021). Encounters in a marginalised subject: The experiential challenges faced by Tasmanian Health and Physical Education teachers. *Australian Journal of Education*, 65(1), 24–40. <https://doi.org/10.1177/0004944120934964>
- Cruzado Saldaña, J. J. (2022). La evaluación formativa en la educación. *Comuni@cción: Revista De Investigación En Comunicación Y Desarrollo*, 13(2), 149–160. <https://doi.org/10.33595/2226-1478.13.2.672>
- Curilem, C., Atilo, A., & Tuillang, Y. (2012). Aplicación del Test Course Navette. *Revista Motricidad Humana*, 16(2), 95–99. <https://www.revistamotricidad.com/wp-content/uploads/2016/03/95-99.pdf>
- Da Costa, F. S., dos Santos, Â. C. C. V., Coelho, F. P. D. S. M., da Silva, P. A. C. M., & de Sousa, J. C. (2024). Gordofobia na escola: relações com a Educação Física a partir de uma análise do Estado do Conhecimento. *Cuadernos de Educación y Desarrollo*, 16(6), e4479–e4479. <https://doi.org/10.55905/cuadv16n6-079>
- Danthon, S., Mascaret, N., & Cury, F. (2019). Test anxiety in physical education: The predictive role of gender, age, and implicit theories of athletic ability. *European Physical Education Review*, 26(1), 128–143. <https://doi.org/10.1177/1356336X19839408>
- De Bruijn, A., Mombarg, R., & Timmermans, A. C. (2021). The importance of satisfying children's basic psychological needs in primary school physical education for PE-motivation, and its relations with fundamental motor and PE-related skills. *Physical Education and Sport Pedagogy*, 27(4), 422–439. <https://doi.org/10.1080/17408989.2021.1906217>
- De León, L. M. (2024). Funcionalidad de la Educación Física en el desarrollo integral de los estudiantes de secundaria. *Delectus*, 7(1), 1–11. <https://doi.org/10.36996/delectus.v7i1.234>

- Del Pino-Sepúlveda, M., & Montanares-Vargas, E. (2019). Evaluación comunicativa y selección de contenidos en contextos escolares vulnerables chilenos. *Revista Electrónica de Investigación Educativa*, 21(3), 1–12. <https://doi.org/10.24320/redie.2019.21.e03.1984>
- Delgado, M. y Llorca, J. (2004). Estudios longitudinales: concepto y particularidades. *Revista Española de Salud Pública*, 78(2), 141–148. https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1135-57272004000200002
- Del Villar Á., F. (1972). La credibilidad de la investigación cualitativa en la enseñanza de la educación física. *Apunts. Educación Física y Deportes*, 37, 26–33.
- Denecke, K., Glauser, R., & Reichenpfader, D. (2023). Assessing the Potential and Risks of AI-Based Tools in Higher Education: Results from an eSurvey and SWOT Analysis. *Trends in Higher Education*, 2(4), 667–688. <https://doi.org/10.3390/higheredu2040039>
- Denzin, N. K., & Lincoln, Y. S. (2011). *The SAGE Handbook of Qualitative Research* (4th ed.). SAGE Publications.
- Devís-Coret, N., & Martos García, D. (2023a). Normatividad corporal, gordofobia y género. *Tándem: Didáctica de la educación física*, (82), 7–12. <https://hdl.handle.net/10550/95921>
- Devís-Coret, N., & Martos-Garcia, D. (2023b). El discurs normatiu de la corporalitat, la ‘grassofòbia’ i el gènere: una proposta provocativa en Educació Física. *Temps d’Educació*, 65, 177–194. <https://doi.org/10.1344/TempsEducacio2023.65.11>
- Devís-Devís, J. (2017). La investigación narrativa en la educación física y el deporte. *Movimiento*, 23(1), 13–24. <https://doi.org/10.22456/1982-8918.71277>
- Díaz-Pinzón, J. E. (2021). Análisis de los resultados de la prueba PISA 2018 en matemáticas para América. *Revista de Investigaciones Universidad del Quindío*, 33(1), 104–114. <https://doi.org/10.33975/riuq.vol33n1.463>
- Dinan-Thompson, M., & Penney, D. (2015). Assessment literacy in primary physical education. *European Physical Education Review*, 21(4), 485–503. <https://doi.org/10.1177/1356336X15584087>

- Dowling, F., Garrett, R., & Wrench, A. (2013). Narrative inquiry in physical education research: The story so far and its future promise. *Sport, Education and Society*, 20(7), 924–940. <https://doi.org/10.1080/13573322.2013.857301>
- Drury, D., & Baer, J. (2011). *The American public school teacher: Past, present, and future*. Harvard Education Press.
- Duque Ramos, V. H., Reina Román, M., Mancha Triguero, D., Ibáñez Godoy, S. J., & Sáenz López, P. (2021). Relación de la carga de entrenamiento con las emociones y el rendimiento en baloncesto formativo. *Retos. Nuevas tendencias en Educación Física, Deporte y Recreación*, 40, 164–173. <https://doi.org/10.47197/retos.vli40.82441>
- Durán, F. (2017). La evaluación de la calidad educativa en Chile: instrumentos de control y rendición de cuentas. *Revista Educación, Política y Sociedad*, 3(1), 85–99. <https://doi.org/10.15366/rep2018.3.1.005>
- Elif, K., & Zekiye, D. E. (2020). Examination of intrinsic motivation, and job satisfaction of physical education and other branch teachers. *Educational Research and Reviews*, 15(10), 632–638. <https://doi.org/10.5897/ERR2020.4040>
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119–161). Macmillan.
- Escalante-Izeta, E. I., Haua-Navarro, K., Moreno-Landa, L. I., & Bertha Pérez-Lizaur, A. (2016). Asociación de la ansiedad y autopercepción corporal con la obesidad infantil Variables nutricias asociadas con la ansiedad y la autopercepción corporal en niñas y niños mexicanos de acuerdo con la presencia de sobrepeso/obesidad. *Salud Mental*, 39(3), 157–163. <https://doi.org/10.17711/SM.0185-3325.2016.014>
- Escudero, J. (2010). Evaluación de las políticas educativas: cuestiones perennes y retos actuales. *Revista Fuentes*, 10, 10–31.
- Essien, A., Bukoye, O. T., O'Dea, X. C., & Kremantzis, M. D. (2024). The influence of AI text generators on critical thinking skills in UK business schools. *Studies in Higher Education*, 49(5). <https://doi.org/10.1080/03075079.2024.2316881>

- Estevan, I., Bardid, F., Utesch, T., Menescardi, C., Barnett, L. M., & Castillo, I. (2021). Examining early adolescents' motivation for physical education: associations with actual and perceived motor competence. *Physical Education and Sport Pedagogy*, 26(4), 359- 374. <https://doi.org/10.1080/17408989.2020.1806995>
- Estupiñán R., J., Batista Hernández, N., Real Zumba, G., Vera Márquez, M. C., & Oviedo Balla, B. W. (2023). El assessment center para la evaluación de las competencias adquiridas por los estudiantes de nivel superior. *Investigación Operacional*, 40(5). <https://revistas.uh.cu/invoperacional/article/view/517>
- European Ministers of Education. (1999). *The Bologna Declaration of 19 June 1999*. <https://www.ehea.info/page-ministerial-declarations>
- Evans, C. (2018). *Analysing semi-structured interviews using thematic analysis: exploring voluntary civic participation among adults*. In J. Lewis (Ed.), SAGE Research Methods Datasets Part 1. SAGE Publications, Ltd. <https://doi.org/10.4135/9781526439284>
- Eversley, G., y González-Moreno, A. (2020). Evaluación auténtica como alternativa para la mejora en el proceso de enseñanza-aprendizaje de educación física en escuelas primarias. *Revista Ciencias de la Actividad Física*, 21(2), 1-18. <https://doi.org/10.29035/rcaf.21.2.4>
- Facione, P. (2007). Pensamiento Crítico: ¿Qué es y por qué es importante? *Insight assessment*, 22, 23-56. <https://eduteka.icesi.edu.co/modulos/6/134/733/1>
- Falabella, A., & Claudio Ramos Zincke. (2019). La larga historia de las evaluaciones nacionales a nivel escolar en Chile. *Cuadernos Chilenos De Historia De La Educación*, (11), 66–98. <https://doi.org/10.60611/cche.vi11.133>
- Fernández-Espínola, C., Abad Robles, M. T., Collado-Mateo, D., Almagro, B. J., Castillo Viera, E., & Giménez Fuentes-Guerra, F. J. (2020). Effects of cooperative-learning interventions on physical education students' intrinsic motivation: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*, 17(12), 44–51. <https://doi.org/10.3390/ijerph17124451>

- Fernández-Ferrer, M., & Forés Miravalles, A. (2016). Evaluación del desarrollo competencial en la educación superior. *Educación*, 54(2), 391. <https://doi.org/10.5565/rev/educar.799>
- Fernández Garcimartín, C. (2023). *Evaluación formativa y compartida en el desarrollo y evaluación de los Trabajos de Fin de Grado de Educación Física: análisis y propuestas*. (Tesis doctoral, Universidad de Valladolid). <https://doi.org/10.35376/10324/66281>
- Fernández-Garcimartín, C., López-Pastor, V. M., Fuentes-Nieto, T., & Hortigüela-Alcalá, D. (2023). Evaluación formativa y compartida para la tutorización de Trabajos Fin de Grado en la Formación del Profesorado de Educación Física. *Cultura, Ciencia y Deporte*, 18(55), 33-56. <https://doi.org/10.12800/ccd.v18i55.1922>
- Ferrer, G. (2006). *Sistemas de evaluación de aprendizajes en América Latina: Balance y desafíos*. Ministerio de Educación del Perú. <https://hdl.handle.net/20.500.12799/214>
- Ferrer, G., & Fiszbein, A. (2015). ¿Qué ha sucedido con los sistemas de evaluación de aprendizajes en América Latina? Lecciones de la última década de experiencia. *Banco Mundial/Diálogo Interamericano*. <https://www.thedialogue.org/wp-content/uploads/2015/09/Spanish-EDU-Learning-Assessment-Systems-v4.pdf>
- Fierro-Suero, S., Almagro, B. J., Sáenz-López, P., & Carmona-Márquez, J. (2020). Perceived novelty support and psychological needs in physical education. *International Journal of Environmental Research and Public Health*, 17, 4169. <https://doi.org/10.3390/ijerph17114169>
- Fierro-Suero, S., Castillo, I., Almagro, B. J., & Sáenz-López, P. (2023). The role of motivation and emotions in physical education: Understanding academic achievement and the intention to be physically active. *Frontiers in Psychology*, 14, Article 1253043. <https://doi.org/10.3389/fpsyg.2023.1253043>
- Finlay, L. (2002). Negotiating the swamp: the opportunity and challenge of reflexivity in research practice. *Qualitative Research*, 2(2), 209–230. <https://doi.org/10.1177/146879410200200205>

- Finol de Franco, M., & Arrieta, X. (2021). Métodos de investigación cualitativa: Un análisis documental. *Encuentro Educacional*, 28(1), 9–28. <https://doi.org/10.5281/zenodo.8169472>
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Pearson Education.
- Fleta, P. G., Serra, M. M., Bahillo, G. R., & Burgués, P. L. (2014). La educación emocional a través de las prácticas motrices de expresión. Perspectiva de género. *Educatio Siglo XXI*, 32(1), 49-70. <https://doi.org/10.6018/j/194081>
- Flores-Aguilar, G., Iniesta-Pizarro, M. & Fernández-Río, J. (2023). “PE Money Heist”: Gamification, Motivational Regulations and Qualifications in Physical Education. *Apunts Educación Física y Deportes*, 151, 36-48. [https://doi.org/10.5672/apunts.2014-0983.es.\(2023/1\).151.04](https://doi.org/10.5672/apunts.2014-0983.es.(2023/1).151.04)
- Foglia, G. E. (2014). Evaluación en educación física: concepciones y prácticas docentes. Voces Y Silencios. *Revista Latinoamericana De Educación*, 5(2), 99-117. <https://doi.org/10.18175/vys5.2.2014.02>
- Fox, C. (2012). How teachers can use PE Metrics for grading. *Journal of Physical Education, Recreation & Dance*, 83(5), 16–34. <https://doi.org/10.1080/07303084.2012.10598774>
- Fraile, A. (2015). La evaluación formativa y cooperativa en Educación Física. *Infancia, Educación y Aprendizaje*, 1(1), 198–223. <https://doi.org/10.22370/ieya.2015.1.1.578>
- Fraile, J., Gil-Izquierdo, M., Zamorano-Sande, D., & Sánchez-Iglesias, I. (2020). Autorregulación del aprendizaje y procesos de evaluación formativa en los trabajos en grupo. *RELIEVE - Revista Electrónica De Investigación Y Evaluación Educativa*, 26(1). <https://doi.org/10.7203/relieve.26.1.17402>
- Fraile, J., Pardo, R., y Panadero, E. (2017). ¿Cómo emplear las rúbricas para implementar una verdadera evaluación formativa? *Revista Complutense de Educación*, 28(4), 1321–1334. <https://doi.org/10.5209/RCED.51915>

- Freire, E. (2021). Importancia de la retroalimentación formativa en el proceso de enseñanza-aprendizaje. *Universidad y Sociedad*, 13(4), 389–397. <https://rus.ucf.edu.cu/index.php/rus/article/view/2178>
- Freire, P. (1970). *Pedagogía del oprimido* (M. A. Ramos, Trans.). Siglo XXI. (Original work published 1970).
- Frossard, M., Stieg, R., Santos, W. (2021). Prácticas evaluativas en tres cursos de educación física en Sudamérica. *Alteridad*, 16. <http://dx.doi.org/10.17163/alt.v16n2.2021.04>
- Fuentes-Diego, V., & Salcines-Talledo, I. (2018). Estudio sobre la Implementación de la Evaluación Formativa y Compartida en un Ciclo Formativo de Grado Superior. *Revista Iberoamericana de Evaluación Educativa*, 11. <https://doi.org/10.15366/riee2018.11.2.005>
- Fuentes, T., & López Pastor, V. (2017). Evaluación auténtica, coevaluación y uso de las TIC en educación física: un estudio de caso en secundaria. *Revista Infancia, Educación y Aprendizaje*, 3(2). <https://doi.org/10.22370/ieya.2017.3.2.697>
- Fuentes-Vásquez, B., Gutiérrez-Verdugo, N., y Gutiérrez-Sepúlveda, K. (2025). Tensión entre la gobernanza neoliberal y la evaluación educativa en Chile: una revisión sistemática. *Revista Actualidades Investigativas en Educación*, 25(1), 1-36. <https://doi.org/10.15517/aie.v25i1.60669>
- Fulano-Vargas, B. C., & Meneses-Runza, I. N. (2023). La evaluación escolar desde la perspectiva de los estudiantes. *RECIE. Revista Caribeña de Investigación Educativa*, 7(1), 163-182. <https://doi.org/10.32541/recie.2023.v7i1.pp163-182>
- Galán, C. H., & López, A. T. (2012). La evaluación formativa y compartida en la formación inicial del profesorado desde la perspectiva del alumnado y de los egresados. *Psychology, Society & Education*, 4(1), 103–116. <https://doi.org/10.25115/psye.v4i1.484>
- Galaz-Ruiz, A., & Anaya-Torres, E. D. (2023). Repensando la evaluación docente: Chile–México. *Estudios Pedagógicos*, 49(1), 203–218. <https://doi.org/10.4067/S0718-07052023000100203>
- Gall, H. (2015). Evaluación en educación física, orientaciones para la práctica. *Educación Física Y Deporte*, 5(3), 4–18. <https://doi.org/10.17533/udea.efyd.23077>

- Gallardo-Fuentes, F. J., Carter-Thuillier, B., & López-Pastor, V. M. (2019). Evaluación Formativa y Compartida en la Formación Inicial del Profesorado de Educación Física: Resultados tras Cuatro Años de Implementación en una Universidad Pública Chilena. *Revista Iberoamericana De Evaluación Educativa*, 12(1). <https://doi.org/10.15366/rie2019.12.1.008>
- Gallardo-Fuentes, F. J., López-Pastor, V. M., & Carter Thuillier, B. (2017). ¿Hay evaluación formativa y compartida en la formación inicial del profesorado en Chile? Percepción de alumnado, profesorado y egresados de una Universidad. *Psychology, Society and Education*, 9(2), 227–238. <https://doi.org/10.25115/psye.v9i2.699>
- Gallardo, F., Carter-Thuilier, B. I., López-Pastor, V. M., Ojeda-Nahuelcura, R., & Fuentes-Nieto, T. (2022). Sistemas de evaluación en la formación del profesorado de Educación Física: un estudio de casos en contexto chileno. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, 43, 117–126. <https://doi.org/10.47197/retos.v43i0.88570>
- García-Gámez, G. (2024). La evaluación como herramienta para mejorar los aprendizajes: la retroalimentación y la evaluación auténtica. *Revista Latinoamericana Ogmios*, 4. <https://doi.org/10.53595/rlo.v4.i9.091>
- García Herranz, S., & López-Pastor, V. (2015). Evaluación formativa y compartida en educación infantil. *Qualitative Research in Education*, 4(1), 22–49. <https://doi.org/10.17583/qre.2015.1269>
- García-Jaén, M. (2020). Gamification in Physical Education: Evaluation of Impact on Motivation and Academic Performance within Higher Education. *International Journal of Environmental Research and Public Health*, 17(12), 4465. <https://doi.org/10.3390/ijerph17124465>
- García, J. y García, M. (2022). La evaluación por competencias en el proceso de formación. *Revista Cubana de Educación Superior*, 41(2). <http://ref.scielo.org/d8hgtz>
- George, E. B., & González-Moreno, P. A. (2020). Evaluación auténtica como alternativa para la mejora en el proceso de enseñanza-aprendizaje de educación física en escuelas primarias. *Revista Ciencias de la Actividad Física*, 21(2), Art. 3. <https://doi.org/10.29035/rcaf.21.2.3>

- Gerstl-Pepin, C., & Patrizio, K. (2009). Learning from Dumbledore's pensieve: Metaphor as an aid in teaching reflexivity in qualitative research. *Qualitative Research*, 9(3), 299–308. <https://doi.org/10.1177/1468794109105026>
- Gil Madrona, P., Cuevas Campos, R., Contreras Jordán, O. R., & Díaz Suárez, A. (2012). Educación Física y hábitos de vida activa: Percepciones de los adolescentes y relación con el abandono deportivo. *Aula Abierta*, 40(3), 115–124. <http://hdl.handle.net/11162/4234>
- Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24(1), 355–392. <https://doi.org/10.3102/0091732X024001355>
- Gobierno de Chile. (2023). *Resolución Exenta N.º 2106: Procedimientos e instrumentos del sistema de acceso a las instituciones de educación superior*. Ministerio de Educación. <https://bcn.cl/39trq>
- Gómez, I. R., del Valle Díaz, M. S., & De la Vega Marcos, R. (2018). Revisión nacional e internacional de las competencias profesionales de los docentes de Educación Física. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, 34, 383–388. <https://doi.org/10.47197/retos.v0i34.58609>
- Gómez, M., Aranda, E., & Santos, J. (2017). A competency model for higher education: An assessment based on placements. *Studies in Higher Education*, 42(12), 2195–2215. <https://doi.org/10.1080/03075079.2016.1138937>
- González, G. (2013). *Evolución de la identidad profesional de un docente novel de Educación Física: Análisis a partir de la reflexión de sus experiencias personales y de su propia práctica* [Tesis doctoral, Universidad de Valladolid]. Repositorio Documental de la Universidad de Valladolid. <https://doi.org/10.35376/10324/2817>
- González-Cutre, D., Jiménez-Loaisa, A., Abós, Á., & Ferriz, R. (2021). *Estrategias motivacionales para incluir novedad y variedad en educación física*. En I. Aguaded (Ed.), *Innovación en educación física y deporte: Nuevas tendencias y propuestas didácticas* (pp. 99–102). Editorial Universidad de Zaragoza. <https://doi.org/10.26754/uz.978-84-18321-22-1>

- González-Cutre, D., & Sicilia, A. (2019). The importance of novelty satisfaction for multiple positive outcomes in physical education. *European Physical Education Review*, 25(3), 859–875. <https://doi.org/10.1177/1356336X18783980>
- Goodson, I. (2013). *Storying the self: Life politics and the study of the teacher's life and work*. In Clandinin, D. J. (Ed.), *Handbook of narrative inquiry* (pp. 273–293). SAGE Publications.
- Greene, J. C. (2007). *Mixed methods in social inquiry*. Jossey-Bass.
- Groves, S., & Laws, C. (2003). The Use of Narrative in Accessing Children's Experiences of Physical Education. *European Journal of Physical Education*, 8(2), 160–174. <https://doi.org/10.1080/1740898030080205>
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Sage.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Sage.
- Guidot, A. G. (2022). *Estigmatización de las corporalidades gordas femeninas: Gordofobia “una herramienta odiológica de dominación”* (Tesis doctoral, Universidad Nacional de Luján). <http://ri.unlu.edu.ar/xmlui/handle/rediunlu/1861>
- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, 52(3), 67–86. <https://doi.org/10.1007/BF02504676>
- Hamodi, C., López Pastor, V. M., & López Pastor, A. T. (2015). Medios, técnicas e instrumentos de evaluación formativa y compartida del aprendizaje en educación superior. *Perfiles Educativos*, 37(147). <https://doi.org/10.22201/iisue.24486167e.2015.147.47271>
- Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365–379. <https://doi.org/10.1080/0969594970040304>
- Harper, D. (2001). *Evaluation*. In Online Etymology Dictionary. <https://www.etymonline.com/word/evaluation>

- Harte, Naomi P., Alfrey, L., Spray, Christopher M. y C. (2023). The if, why and how of fitness testing in secondary school physical education. *European Physical Education Review*, 29(3), 475–492. <https://doi.org/10.1177/1356336X231219937>
- Hay, P. J., & Penney, D. (2009). Proposing conditions for assessment efficacy in physical education. *European Physical Education Review*, 15(3), 389-405. <https://doi.org/10.1177/1356336X09364294>
- Hernández, R. M. (2014). La investigación cualitativa a través de entrevistas: su análisis mediante la teoría fundamentada. *Cuestiones Pedagógicas*, 23, 187-210. <http://hdl.handle.net/11441/36261>
- Hernández, C. I. (2020). La perspectiva de género en las prácticas deportivas: hacia una visión coeducativa en educación física. *Diálogos sobre Educación. Temas Actuales en Investigación Educativa*, 11(20). <https://doi.org/10.32870/dse.v0i20.592>
- Herrera, I. R. (2018). Evaluación para el aprendizaje. *Revista Educación las Américas*, 6, 13-28. <https://revistas.udla.cl/index.php/rea/article/view/22>
- Herrero-González, D., López-Pastor, V. M., Manrique-Arribas, J. C., & Moura, A. (2023). Formative and shared assessment: Literature review on the main contributions in physical education and physical education teacher education. *European Physical Education Review*, 0(0). <https://doi.org/10.1177/1356336X231220995>
- Herrero-González, D., Manrique Arribas, J. C., & López-Pastor, V. M. (2021). Incidencia de la formación inicial y permanente del profesorado en la aplicación de la evaluación formativa y compartida en Educación Física. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, 41, 533–543. <https://doi.org/10.47197/retos.v0i41.86090>
- Hodge, K., & Sharp, L.-A. (2016). Case studies. In B. Smith & A. C. Sparkes (Eds.), *Routledge handbook of qualitative research in sport and exercise* (pp. 62–74). Routledge. <https://doi.org/10.4324/9781315762012>
- Hortigüela-Alcalá, D., González-Víllora, S., & Hernando-Garijo, A. (2021). Do we really assess learning in physical education? Teachers' perceptions at different educational stages. *Retos:*

- Nuevas Tendencias en Educación Física, Deporte y Recreación*, 42, 643–654.
<https://doi.org/10.47197/retos.v42i0.88686>
- Hortigüela-Alcalá, D., Pérez-Pueyo, Á., & del Río, J. (2016). Influencia de las experiencias vivenciadas por el alumnado en el desempeño de futuras competencias docentes. *Contextos Educativos: Revista de Educación*, (19), 25–41. <https://doi.org/10.18172/con.2651>
- Hortigüela Alcalá, D., Pérez Pueyo, Á., & González Calvo, G. (2019). Pero... ¿a qué nos referimos realmente con la evaluación formativa y compartida? Confusiones habituales y reflexiones prácticas. *Revista Iberoamericana de Evaluación Educativa*, 12(1), 13–27.
<https://doi.org/10.15366/riee2019.12.1.001>
- Hortigüela-Alcalá, D., Pérez-Pueyo, A., & Salicetti, A. (2015). ¿Cómo percibe el alumnado universitario de educación física la evaluación recibida? Contraste de dos metodologías diferentes. *Retos. Nuevas tendencias en Educación Física, Deporte y Recreación*, 28, 66-70.
<https://doi.org/10.47197/retos.v0i28.34845>
- Hugo, V., Ramos, D., Román, M. R., Triguero, D. M., José, S., Godoy, I., & Buñuel, P. S. (2021). Relación de la carga de entrenamiento con las emociones y el rendimiento en baloncesto formativo. *Retos*, 40, 164–173. <https://doi.org/10.47197/retos.v1i40.82441>
- Huisman, B., Saab, N., Van Den Broek, P., & Van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: a Meta-Analysis. *Assessment & Evaluation in Higher Education*, 44(6), 863-880.
<https://doi.org/10.1080/02602938.2018.1545896>
- Hung, N. M., Thuy, V. X., Dung, N. V., Khuong, T. M., Hoa, L. T. P., Linh, L. T., & Thuc, D. C. (2024). The Evaluation Criteria for Graduating Students in Physical Education Majors are Based on a Competency-Based Approach. *Revista De Gestão Social E Ambiental*, 18(4), e07422. <https://doi.org/10.24857/rgsa.v18n4-152>
- Hurtado-Guerrero, M., Zavala-Crichton, J. P., del Val Martín, P., Báez-San Martín, E., Yáñez-Sepúlveda, R., & Hinojosa-Torres, C. (2024). La evaluación en educación física: Representaciones sociales de profesores de establecimientos educacionales de Viña del Mar,

- Chile. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 59, 934–943. <https://doi.org/10.47197/retos.v59.107651>
- James, A. R. (2018). Grading in physical education. *Journal of Physical Education, Recreation & Dance*, 89(5), 5–7. <https://doi.org/10.1080/07303084.2018.1442063>
- Jiménez Barret, I. (2017). Una experiencia de Evaluación Formativa en Educación Física en Educación Secundaria. *Revista Infancia, Educación y Aprendizaje*, 3. <http://dx.doi.org/10.22370/ieya.2017.3.2.708>
- Jiménez, S., & González-Palomares, A. (2023). “ODS 5. Igualdad de género” y Educación Física: propuesta de intervención mediante los deportes alternativos. *Retos*, 49, 595–602. <https://doi.org/10.47197/retos.v49.95791>
- Junior, M. de P. P., Ramírez, C. A. N., Gutiérrez, P. A. P., Espinoza, G. A. R., & Retamal, S. E. V. (2022). Contenidos curriculares, aprendizajes esperados e instrumentos de evaluación utilizados por los docentes de Educación Física en Enseñanza Media de la comuna de Talca: un estudio piloto. *Journal of Movement & Health*, 19(2). [https://doi.org/10.5027/jmh-Vol19-Issue2\(2022\)art150](https://doi.org/10.5027/jmh-Vol19-Issue2(2022)art150)
- Kalajas-Tilga, H., Koka, A., Hein, V., Tilga, H., & Raudsepp, L. (2020). Motivational processes in physical education and objectively measured physical activity among adolescents. *Journal of Sport and Health Science*, 9(5), 462–471. <https://doi.org/10.1016/j.jshs.2019.06.001>
- Kasneci, E., Sessler, K., Kühn, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Killian, C. M., & Woods, A. M. (2021). Assessment practices in K–12 physical education in the United States. *Research Quarterly for Exercise and Sport*, 92(3), 248–258. <https://doi.org/10.1080/02701367.2021.1894315>
- Kincheloe, J. L. (2001). Describing the Bricolage: Conceptualizing a New Rigor in Qualitative Research. *Qualitative Inquiry*, 7(6), 679–692. <https://doi.org/10.1177/107780040100700601>

- Kirk, D. (1990). *Educación física y currículum* (Vol. 3). Universitat de València
- Kirk, D. (2009). Physical Education Futures. *Perlego*, 1–184. <https://doi.org/10.4324/9780203874622>
- Kleinman, I. (1997). Grading: A powerful teaching tool. *Journal of Physical Education, Recreation & Dance*, 68(5), 29–32. <https://doi.org/10.1080/07303084.1997.10604942>
- Krijgsman, C., Vansteenkiste, M., Van Tartwijk, J., Maes, J., Borghouts, L., Cardon, G., Mainhard, T., & Haerens, L. (2017). Performance grading and motivational functioning and fear in physical education: A self-determination theory perspective. *Learning and Individual Differences*, 55, 202–211. <https://doi.org/10.1016/j.lindif.2017.03.017>
- Kvale, S., & Brinkmann, S. (2009). *InterViews: Learning the craft of qualitative research interviewing* (2nd ed.). SAGE Publications.
- Landi, D. (2023). Thinking Qualitatively: Paradigms and Design in Qualitative Research. En Richards, K.A., M.A. Hemphill, Wright, P.M. (eds.). *Qualitative Research and Evaluation in Physical Education and Sport Pedagogy*, 31-47. Jones & Bartlett Learnings.
- Lee, H. (2023). The rise of ChatGPT: Exploring its potential in medical education. *Anatomical sciences education, In Press*. <https://doi.org/10.1002/ase.2270>
- Leenknecht, M., Wijnia, L., Köhlen, M., Fryer, L., Rikers, R., & Loyens, S. (2021). Formative assessment as practice: The role of students’ motivation. *Assessment & Evaluation in Higher Education*, 46(2), 236–255. <https://doi.org/10.1080/02602938.2020.1765228>
- Lentillon-Kaestner, V. & Voisard, N. (2023). Pratiques évaluatives en EPS dans trois contextes avec un programme commun et des modes de communication de l’évaluation sommative différents: les apports d’une approche mixte. *Staps*, 141, 129-144. <https://doi.org/10.3917/sta.141.0129>
- León León, Giselle. (2015). *La narrativa, como recurso en la investigación educativa*. Praxis Investigativa ReDIE, 7(13), 85–92. <https://biblat.unam.mx/es/buscar/la-narrativa-como-recurso-en-la-investigacion-educativa>

- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Sage Publications.
- Lituma Pizarro, J., Aldas Arcos, H. (2023). Metas de logro, percepción de éxito, diversión y persistencia-esfuerzo en la educación física escolar. *Explorador Digital*, 7. <http://dx.doi.org/10.33262/exploradordigital.v7i4.1.2755>
- Lizandra, J., Valencia-Peris, A., Gago, R. A., & Martos-García, D. (2017). Itinerarios de evaluación y su relación con el rendimiento académico. *REDU. Revista de Docencia Universitaria*, 15(2), 315-328. <https://doi.org/10.4995/redu.2017.7862>
- Lombardi, M. M. (2007). Authentic learning for the 21st century: An overview. *Educause Learning Initiative*, 1-12. <https://www.sciepub.com/reference/290496>
- López Ávila, C. R., Arcila-Rodríguez, W. O., & Betancur Agudelo, J. E. (2022). Prácticas evaluativas en la clase de Educación Física: Un asunto de desconocimiento. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 44, 77–86. <https://doi.org/10.47197/retos.v44i0.88630>
- López, A., y Osorio, K. (2016). Percepciones de estudiantes sobre la retroalimentación formativa en el proceso de evaluación. *Actualidades Pedagógicas*, 1(68), 43. <https://doi.org/10.19052/ap.2829>
- López-Gil, J. F., Camargo, E., & Lucas, J. L. Y. (2020). Capacidad aeróbica en escolares de Educación Primaria determinada mediante el test Course Navette: una revisión sistemática. *Revista Electrónica Interuniversitaria de Formación Del Profesorado*, 23(3), 217–232. <https://doi.org/10.6018/REIFOP.402151>
- López Pastor, V. M., Monjas Aguado, R., Gómez García, J., López Pastor, E., Martín Pinela, J., González Badiola, J., Barba Martín, J. J., Aguilar Baeza, R., González Pascual, M., Heras Bernardino, C., Martín, M. I., Manrique Arribas, J. C., & Marugán García, L. (2006). La evaluación en educación física. Revisión de los modelos tradicionales y planteamiento de una alternativa: La Evaluación Formativa y Compartida. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 10, 31–40. <https://doi.org/10.47197/retos.v0i10.35061>

- López-Pastor, V. (2000). Buscando una evaluación formativa en educación física: Análisis crítico de la realidad existente, presentación de una propuesta y análisis general de su puesta en práctica. *Apunts. Educación Física y Deportes*, 4(62), 16-26.
- López Pastor, V. M., Monjas Aguado, R., Gómez García, J., López Pastor, E., Martín Pinela, J., González Badiola, J., Barba Martín, J. J., Aguilar Baeza, R., González Pascual, M., Heras Bernardino, C., Martín, M. I., Manrique Arribas, J. C., & Marugán García, L. (2006). La evaluación en educación física. Revisión de los modelos tradicionales y planteamiento de una alternativa: La Evaluación Formativa y Compartida. *Retos*, 10, 31–40. <https://doi.org/10.47197/retos.v0i10.35061>
- López Pastor, V. M., & García Herranz, S. (2017). Una experiencia de estimulación temprana y evaluación formativa en educación infantil. En V. M. López Pastor & Á. Pérez Pueyo (Coords.), *Evaluación formativa y compartida en educación: Experiencias de éxito en todas las etapas educativas* (pp. 119–125). Universidad de León. <https://uvadoc.uva.es/handle/10324/47605>
- López-Pastor, V. (2011). El papel de la evaluación formativa en la evaluación por competencias: aportaciones de la red de evaluación formativa y compartida en docencia universitaria. *REDU. Revista de Docencia Universitaria*, 9. <https://doi.org/10.4995/redu.2011.6185>
- López-Pastor, V. M. (2009). Evaluación formativa y compartida en educación superior: propuestas, técnicas, instrumentos y experiencias. Madrid: Narcea, 270 páginas. *Bordón. Revista De Pedagogía*, 62(1), 169–170. <https://recyt.fecyt.es/index.php/BORDON/article/view/29114>
- López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2012). Alternative assessment in physical education: A review of international literature. *Sport, Education and Society*, 18(1), 57–76. <https://doi.org/10.1080/13573322.2012.713860>
- López-Pastor, V., & Sicilia-Camacho, A. (2017). Formative and shared assessment in higher education. Lessons learned and challenges for the future. *Assessment & Evaluation in Higher Education*, 42(1), 77-97. <https://doi.org/10.1080/02602938.2015.1083535>
- López-Pastor, V. M., & Pérez-Pueyo, Á. (2024). Evaluación de la condición física: Razones para superar el modelo tradicional y revisión de alternativas educativas. En Á. Pérez-Pueyo, D.

- Hortigüela Alcalá, C. Gutiérrez García, & R. A. Barba Martín (Coords.), *Evaluación formativa y compartida en Educación Física: Fundamentos y experiencias prácticas en todas las etapas educativas* (pp. 97–124). Universidad de León, Servicio de Publicaciones. <https://hdl.handle.net/10612/22962>
- López, R. C., Chihuailaf-Vera, M. L., Lopicich, A. P., & Moraga, A. C. (2022). Percepciones del profesorado chileno de educación física hacia la perspectiva de género. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 45, 786–795. <https://doi.org/10.47197/retos.v45i0.91022>
- Lozano, S. J., & Palomares, A. G. (2023). ODS 5. Igualdad de género y Educación Física: propuesta de intervención mediante los deportes alternativos. *Retos: nuevas tendencias en educación física, deporte y recreación*, 49, 595-602. <https://doi.org/10.47197/retos.v49.95791>
- Luarde Rocha, C., Garrido Méndez, A., Pacheco Carrillo, J., & Daolio, J. (2016). Antecedentes históricos de la actividad física para la salud. *Revista Ciencias de la Actividad Física*, 17(1), 67–76. Recuperado de <https://www.redalyc.org/articulo.oa?id=525664802007>
- Llapur, E. (2023). La evaluación en la Educación Física: una problemática de la educación contemporánea. *Educación Física y Ciencia*, 25, e255. <https://www.redalyc.org/journal/4399/439974899003/html/>
- Lleixà Arribas, T., Soler i Prat, S., & Serra, P. (2020). Perspectiva de género en la formación de maestras y maestros de Educación Física. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 37, 634–642. <https://doi.org/10.47197/retos.v37i37.74253>
- Lleixà, T., Torralba, M., y Abrahão, S. (2010). Evaluación de competencias en Educación Física: Investigación-acción para el diseño de procedimientos de evaluación en la Etapa Primaria. *Movimiento*, 16(4), 33-51. <https://www.redalyc.org/articulo.oa?id=115316963003>
- Llergo Young, M., Ambris Sandoval, J. C., Mato Medina, O. E., & Matos Ceballos, J. J. (2021). Motivación en la práctica de Educación Física, en adolescentes que cursan el nivel Medio Superior. *Revista Metropolitana de Ciencias Aplicadas*, 4(2), 260-267. <https://doi.org/10.62452/srkja846>

- Mackley, H., Edwards, S., McLean, K., & Cinelli, R. (2022). Building collaborative competencies through play with outdoor loose parts materials in primary school. *International Journal of Environmental Research and Public Health*. <https://doi.org/10.1080/0305764X.2022.2030300>
- Madaus, G. F., Scriven, M., & Stufflebeam, D. L. (Eds.). (2012). *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed.). Evaluation in education and human services. <https://scholarworks.wmich.edu/books/514>
- Mantilla-Falcón, L. M., Miranda Ramos, D. P., Ortega Zurita, G. E., & Meléndez-Tamayo, C. F. (2020). Hibridación de modelos pedagógicos en la práctica docente en la educación superior en Ecuador. Caso Universidad Técnica de Ambato. *Cuadernos de Investigación Educativa*, 11(1), 85–101. <https://doi.org/10.18861/cied.2020.11.1.2944>
- Mao, Y., Nguyen, D., & Zawacki-Richter, O. (2024). Adaptive feedback systems using AI: A review. *Educational Technology Research and Development*, 72(1), 1–20. <https://doi.org/10.1007/s11423-023-10212-5>
- Martínez Abajo, J., & Bastarrica Varela, O. (2023). El role-playing de la gordofobia. *Tándem: didáctica de la educación física*. <https://hdl.handle.net/11162/258265>
- Martínez, L. (2007). La observación y el diario de campo en la definición de un tema de investigación. *Revista perfiles libertadores*, 4(80), 73-80.
- Martínez Muñoz, L. F., Pastor, M. L. S., & Oliva, F. J. C. (2017). Percepciones de alumnado y profesorado en Educación Superior sobre la evaluación en formación inicial en educación física. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 32(32), 76–81. <https://doi.org/10.47197/retos.v0i32.52918>
- Martínez Navarro, A. (2010). Anotaciones a la historia de la educación física española en el siglo XIX. *Historia de la Educación*, 2. <https://revistas.usal.es/tres/index.php/0212-0267/article/view/6476>

- Martos García, D., & García Puchades, W. (2023). La evaluación participativa y la negociación con el alumnado: una propuesta basada en la Pedagogía de la Ignorancia. *REDU. Revista de Docencia Universitaria*, 21(1), 129-146. <https://doi.org/10.4995/redu.2023.17372>
- Mastagli, M., Malini, D., Hainaut, J.-P., Van Hoye, A., & Bolmont, B. (2020). Summative assessment versus formative assessment: An ecological study of physical education by analyzing state-anxiety and shot-put performance among French high school students. *Journal of Physical Education and Sport*, 20(Suppl 3), 3298–3305. <https://doi.org/10.7752/jpes.2020.s3298>
- Matanin, M., y Tannehill, D. (1994). Assessment and Grading in Physical Education. *Journal of Teaching in Physical Education*, 13(4), 395–405. <https://doi.org/10.1123/jtpe.13.4.395>
- Matus-Castillo, C., Cornejo-Améstica, M. & Castillo-Retamal, F. (2021). La perspectiva de género en la formación inicial docente en la educación física chilena. *Retos*, 40, 326–335. <https://doi.org/10.47197/retos.v1i40.83082>
- Matus-Castillo, C., Serra, P., Soler, S., Vilanova, A., Flores-Rivera, C., Knijnik, J., & Luna-Villouta, P. (2023). Creencias y prácticas sobre la perspectiva de género en el profesorado de Pedagogía en Educación Física en Chile. *Retos*, 47, 969–977. <https://doi.org/10.47197/retos.v47.96869>
- McArthur, J. (2015). Assessment for social justice: The role of assessment in achieving social justice. *Assessment & Evaluation in Higher Education*, 41(7), 967–981. <https://doi.org/10.1080/02602938.2015.1053429>
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). SAGE Publications.
- Medina Martínez, R., & Díaz Castillo, N. (2022). La aplicación de la evaluación auténtica en sistemas y su impacto en la calidad educativa. *Presencia Universitaria*, 9(18-1). <https://doi.org/10.29105/pu9.18-1>
- Mercier, K., Simonton, K., Centeio, E., Barcelona, J., & Garn, A. (2023). Middle school students' attitudes toward physical activity and physical education, intentions, and physical activity

- behavior. *European Physical Education Review*, 29(1), 40-54.
<https://doi.org/10.1177/1356336X221106051>
- Merellano-Navarro, E., Almonacid-Fierro, A., & Oyarce, M. F. M. (2019). Renaming the pedagogical knowledge: A look from the teaching practice. *Educacao e Pesquisa*, 45(May), 1–18. <https://doi.org/10.1590/s1678-4634201945192146>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, 13(9), 856.
<https://doi.org/10.3390/educsci13090856>
- Ministerio de Educación de Chile. (2010). *Resultados SIMCE 2010. Segundo año medio: Educación Física*. Agencia de Calidad de la Educación.
https://archivos.agenciaeducacion.cl/biblioteca_digital_historica/resultados/2010/result8b_e_dfisica_2010.pdf
- Ministerio de Educación. (2014). *Estándares orientadores para carreras de Pedagogía en Educación Física*. Santiago de Chile: CPEIP. <https://estandaresdocentes.mineduc.cl/wp-content/uploads/2021/08/Educacion-Fisica-2014.pdf>
- Ministerio de Educación. (2016). *Ley N.º 20.903 Crea el Sistema de Desarrollo Profesional Docente y modifica otras normas*. Gobierno de Chile.
<https://www.bcn.cl/leychile/navegar?idNorma=1087343>
- Ministerio de Educación. (2015). *Bases curriculares de Séptimo a Segundo Medio*.
https://www.curriculumnacional.cl/614/articles-37136_bases.pdf
- Ministerio de Educación de Chile – Equipo de Tarea SIMCE. (2015). *Hacia un sistema completo y equilibrado de evaluación de los aprendizajes: Informe ejecutivo*. Santiago, Chile. Recuperado de <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/17445>
- Ministerio de Educación. (2016). *Educación Física y Salud- Programa de estudio Séptimo Básico*.
https://www.curriculumnacional.cl/614/articles-20744_programa.pdf

- Ministerio de Educación de Chile. (2018). *Política para el fortalecimiento de la evaluación en aula*. Santiago, Chile. <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/2255>
- Ministerio de Educación. (2018). Decreto N.º 67 de 2 de agosto de 2018. *Aprueba normas mínimas nacionales sobre evaluación, calificación y promoción y deroga los decretos exentos N.º 511 de 1997, N.º 112 de 1999 y N.º 83 de 2001*, todos del Ministerio de Educación. Diario Oficial de la República de Chile. <https://www.bcn.cl/leychile/navegar?idNorma=1127255>
- Ministerio de Educación. (2021). *Estándares orientadores para carreras de Pedagogía en Educación Física*. Santiago de Chile: CPEIP. <https://estandaresdocentes.mineduc.cl/wp-content/uploads/2021/08/Educacion-Fisica.pdf>
- Ministerio de Educación de Chile. (2023). *Resolución exenta N° 2106: Procedimientos e instrumentos del sistema de acceso a las instituciones de educación superior*. Biblioteca del Congreso Nacional de Chile. <https://bcn.cl/39trq>
- Modell, N., & Gerdin, G. (2022). ‘But in PEH it still feels extra unfair’: Students’ experiences of equitable assessment and grading practices in physical education and health (PEH). *Sport, Education and Society*, 27(9), 1047–1060. <https://doi.org/10.1080/13573322.2021.1965565>
- Molina, J. & Beltrán, V. (2007). Incompetencia motriz e ideología del rendimiento en educación física: el caso de un alumno con discapacidad intelectual. *Motricidad. European Journal of Human Movement*, 19, 165–190. <https://www.eurjhm.com/index.php/eurjhm/article/view/193>
- Molina, M., Pascual-Arias, C., & López-Pastor, V. M. (2020). El rendimiento académico y la evaluación formativa y compartida en formación del profesorado. *ALTERIDAD. Revista de Educación*. <https://doi.org/10.17163/alt.v15n2.2020.05>
- Molina Soria, M., Pascual Arias, C., & López Pastor, V. M. (2020). El uso de sistemas de evaluación formativa y compartida en las aulas de educación física en educación primaria. *Educación Física Y Deporte*, 39(1), 47–71. <https://doi.org/10.17533/udea.efyd.v39n1a04>

- Monforte, J., & Colomer, J. Ú. (2019). ‘Como una chica’: un estudio provocativo sobre estereotipos de género en educación física. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, (36), 74-79. <https://doi.org/10.47197/retos.v36i36.68598>
- Monforte, J., & Pérez-Samaniego, V. (2017). El miedo en Educación Física: una historia reconocible. *Movimiento*, 85-100. <https://doi.org/10.22456/1982-8918.71272>
- Mora, A. I. (2004). La evaluación educativa: Concepto, períodos y modelos. *Actualidades Investigativas en Educación*, 4(2), 1–21. <https://www.redalyc.org/articulo.oa?id=44740211>
- Moral Santaella, C., (2016). Estrategias para resistir a la crisis de confianza en la investigación cualitativa actual. *Educación XXI*, 19(1), 159-177. <https://www.redalyc.org/comocitar.oa?id=70643085007>
- Moreno, A., & Poblete, C. (2015). La educación física chilena y su profesorado: proponiendo algunos retos para la investigación en el área. *Retos*, 28, 291–296. <https://doi.org/10.47197/retos.v0i28.35651>
- Moreno, A., Rivera García, E., & Trigueros Cervantes, C. (2014). Sistema de medición de la calidad de la Educación Física chilena: Un análisis crítico. *Movimiento*, 20(1), 145–167. <http://hdl.handle.net/10481/32996>
- Moreno-Doña, A., Valencia-Peris, A., & Rivera-García, E. (2016). School Physical Education in Three Chilean Schools: a Characterization of their Teaching Practices. *Qualitative Research in Education*, 5(3), 255–275. <https://doi.org/10.17583/qre.2016.2105>
- Moreno Olivos, T. (2016). Evaluación del aprendizaje y para el aprendizaje: Reinventar la evaluación en el aula. Universidad Autónoma Metropolitana-Cuajimalpa. ISBN 978-607-28-0779-2
- Mujica Johnson, F. (2019). Análisis crítico de la formación actitudinal en la asignatura de Educación Física y Salud en Chile. *Revista de Estudios y Experiencias en Educación*, 18, 151–166. <https://www.rexe.cl/index.php/rexe/article/view/785>

- Mújica-Johnson, F.-N., Orellana-Arduiz, N. del C., & Concha-López, R.-F. (2017). Emociones en la clase de Educación Física: revisión narrativa (2010-2016). *Ágora Para La Educación Física y El Deporte*, 19(1), 119. <https://doi.org/10.24197/aefd.1.2017.119-134>
- Mujica Johnson, F. N., Santander Reveco, I., Gajardo Cáceres, P., Concha López, R., Osorio González, A., & Orellana Arduiz, N. (2023). Tensiones epistemológicas en Educación Física. El caso de la exclusión conceptual en los estándares disciplinares para la formación docente en Chile. *Retos*, 49, 225–230. <https://doi.org/10.47197/retos.v49.96419>
- Muñoz, L. F. M., Pastor, M. L. S., & Oliva, F. J. C. (2017). Percepciones de alumnado y profesorado en Educación Superior sobre la evaluación en formación inicial en educación física. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 32(32), 76–81. <https://doi.org/10.47197/retos.v0i32.52918>
- Muñoz-Sepúlveda, F., & Valencia-Peris, A. (2022). Percepción del estudiantado universitario de una experiencia de evaluación basada en competencias: La simulación pedagógica. En I. Moya Mata, E. López-Cañada, & R. Atienza (Eds.), *Libro de Actas del XIV Congreso Nacional en Evaluación Formativa y Compartida en Educación: Tendiendo puentes entre la escuela y la universidad* (pp. 133–135). Publicacions de la Universitat de València. <https://omp.uv.es/index.php/PUV/catalog/book/438>
- Muñoz-Sepúlveda, F., Valencia-Peris, A., Martos-García, D. y Moreno-Doña, A. (en prensa). Dime cuánto corres y te diré qué nota mereces: evaluación en la Educación Física chilena. *Cultura, Ciencia y Deporte*.
- Muñoz-Sepúlveda, F., Vásquez-Cerda, P., Añazco Martínez, L. A., & Molina-Elizalde, J. P. (2025). Aprendizaje y evaluación de la gimnasia artística en la clase de Educación Física: una perspectiva basada en las experiencias del estudiantado de pedagogía. *Sportis. Scientific Journal of School Sport, Physical Education and Psychomotricity*, 11(2), 1–25. <https://doi.org/10.17979/sportis.2025.11.2.11348>
- Muñoz-Sepúlveda, F., Orellana-Fernández, R., & Escobar-Lastra, C. (2025b). “Hombres por aquí, mujeres por allá”: percepciones del profesorado y alumnado sobre la separación por género

- en la clase de Educación Física. *Retos*, 68, 525–539. <https://doi.org/10.47197/retos.v68.116092>
- Muñoz-Sepúlveda, F., Vázquez-Mancilla, B., Acuña-González, H., Sagal-Garrido, G., & Vázquez-Cerda, P. (2024). Práctica temprana en una escuela de educación especial: percepción de competencias inclusivas desde la Educación Física. *Revista de Inclusión Educativa y Diversidad (RIED)*, 2(2), 1–13. <https://doi.org/10.5281/zenodo.13843452>
- Murillo, F. J., & Román, M. (2010). Retos en la evaluación de la calidad de la educación en América Latina. *Revista Iberoamericana de Educación*, 53, 39–60. <https://doi.org/10.35362/rie530559>
- Nabaskués, I., Usabiaga, O., Martos-García, D. & Standal, O. (2019). Comprensión de la habilidad desde la perspectiva del futuro profesorado de Educación Física. *Retos: nuevas tendencias en educación física, deporte y recreación*, (36), 121-128. <https://doi.org/10.47197/retos.v36i36.67598>
- Nadeem, M., Oroszlanyova, M., & Farag, W. (2023). Effect of Digital Game-Based Learning on Student Engagement and Motivation. *Computers*, 12(9), 177. <https://doi.org/10.3390/computers12090177>
- Navarro-Patón, R., Lago-Ballesteros, J., & Arufe-Giráldez, V. (2020). Midiendo la motivación auto-determinada hacia la educación física en la escolaridad obligatoria. *Revista de Psicología del Deporte*, 29(4), 126–134. <http://hdl.handle.net/2183/27566>
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), 149–170. <https://doi.org/10.1080/09695940701478321>
- Nitko, A. J., & Brookhart, S. M. (2013). *Educational assessment of students* (6th ed.). Pearson Education.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, 16(1), 1–13. <https://doi.org/10.1177/1609406917733847>

- Olague, M. Z., & Cañadas, L. (2021). Finalidades de los procesos de evaluación y calificación en Educación Física. Espiral. *Cuadernos del profesorado*, 14(29). <https://doi.org/10.25115/ECP.V14I29.4398>
- Oliver-Álvarez, M., & Martos-García, D. (2023). La evaluación tradicional y sus consecuencias. Un caso en la Educación Física hegemónica. *Cultura, Ciencia y Deporte*, 18(58), 27-49. <https://doi.org/10.12800/ccd.v18i58.1976>
- Oltra, J. A. B. (2017). Aprendiendo a través de tareas de evaluación auténticas. *Revista de Educación*, 376, 25–49. <https://dialnet.unirioja.es/servlet/articulo?codigo=7408500>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2005). *Declaración Universal sobre Bioética y Derechos Humanos*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000146180>
- Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2017). *Educación en Chile: Revisión de políticas nacionales de educación*. París: OECD Publishing. <https://doi.org/10.1787/9789264288720-es>
- Orozco-Jutorán, M. (2006). *La evaluación diagnóstica, formativa y sumativa en la enseñanza de la traducción*. En María José Varela (Ed.), *La evaluación en los estudios de traducción e interpretación* (pp. 47–68). Bienza.
- Ossa, C., & Willatt, C. (2023). Uso de inteligencia artificial generativa para retroalimentar escritura académica en procesos de formación inicial docente. *European Journal of Education and Psychology*, 16(2), 1–16. <https://doi.org/10.32457/ejep.v16i2.2412>
- Otero-Saborido, F. M., Torreblanca-Martínez, V., & González-Jurado, J. A. (2021). Systematic review of self-assessment in physical education. *International Journal of Environmental Research and Public Health*, 18(2), 766. <https://doi.org/10.3390/ijerph18020766>
- Ovalle, A., & Briones Molina, D. N. (2022). Educación física, nacionalismo y eugenesia: El club de gimnasia científica, Chile (1924-1929). *Revista Páginas*, 15(37). <https://doi.org/10.35305/rp.v15i37.721>

- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(5), 533-544. <https://doi.org/10.1007/s10488-013-0528-y>
- Pasek-de Pinto, E. & Mejía, M. T. (2017). Proceso General para la Evaluación Formativa del Aprendizaje. *Revista Iberoamericana de Evaluación Educativa*, 10(1), 177-193. <https://doi.org/10.15366/rie2017.10.1.009>
- Pastor-Vicedo, J., Sánchez-Oliva, A., Sánchez-Blanchart, J., & Martínez-Martínez, J. (2019). Estereotipos de género en educación física. *SPORT TK-Revista EuroAmericana de Ciencias del Deporte*, 8(2), 23–31. <https://doi.org/10.6018/sportk.401071>
- Pavlova, I., Petrytsa, P., Andres, A., Khurtenko, O., Osip, N., Yednak, V., Naumchuk, V., & Mashtaler, I. (2020). Assessment of Student's Competence in Physical Education: Approaches and Methodology. *Revista Romaneasca Pentru Educatie Multidimensionala*, 12(4), 338-356. <https://doi.org/10.18662/rrem/12.4/349>
- Payeras, P. S., Grau, M. P., da Silva Nicolino, A., Prat, S. S., & Silva, A. M. (2020). Género y currículum de formación del profesorado en Educación Física: un diálogo entre Brasil y España. *Revista Iberoamericana de Educación*, 82(2), 191-210. <https://doi.org/10.35362/rie8223637>
- Peiró Velert, C., & Julián Clemente, J. A. (2015). Los modelos pedagógicos en educación física: un enfoque más allá de los contenidos curriculares. *Tándem: Didáctica de la Educación Física*, (50), 69–76. <https://producciocientifica.uv.es/documentos/5eb09ec4299952764113c404>
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). Knowing what students know: The science and design of educational assessment. *National Academy Press*. <https://doi.org/10.17226/10019>
- Peña Troncoso, S., Toro Arevalo, S., Beltrán Veliz, J. C., & Navarro Aburto, B. (2019). Hacia una comprensión de la evaluación conceptual en educación física. *Educación Física y Ciencia*, 21(1), e069. <https://www.redalyc.org/articulo.oa?id=439959077004>

- Pereira, Z. (2011). Los diseños de método mixto en la investigación en educación: una experiencia concreta. *Revista Electrónica Educare*, 15(1), 15–29. <https://doi.org/10.15359/ree.15-1.2>
- Pérez-Pueyo, Á. (2013). El estilo actitudinal: una propuesta para todos y todas desde la inclusión en la educación física. *Lúdica Pedagógica*, 1. <http://dx.doi.org/10.17227/01214128.18ludica81.92>
- Pérez-Pueyo, Á., Hortigüela-Alcalá, D., & Fernandez-Río, J. (2020). Evaluación formativa y modelos pedagógicos: Estilo actitudinal, aprendizaje cooperativo, modelo comprensivo y educación deportiva. *Revista española de educación física y deportes*, 428, 47-66.
- Pérez-Pueyo, Á., Hortigüela Alcalá, D., Herrán Álvarez, I., Vega Cobo, D., Heras Bernardino, C., Garrote García, J., Sobejano Carrocera, M., & Hernando Garijo, A. (2017). La hibridación de modelos pedagógicos en educación física y la evaluación formativa. *Revista Infancia, Educación y Aprendizaje*, 3(2), 411–418. <https://doi.org/10.22370/IEYA.2017.3.2.757>
- Pérez Pueyo, Á. (2012). Las competencias básicas en EF. ¿evaluación o calificación?. *Revista Española De Educación Física Y Deportes*, (397), Pag 37. <https://www.reefd.es/index.php/reefd/article/view/183>
- Pérez Pueyo, Á., Tabernero Sánchez, B., López Pastor, V. M., Ureña Ortín, N., Ruiz Lara, E., Capllonch Bujosa, M., González Fernández, N., & Castejón Oliva, F. J. (2008). Evaluación formativa y compartida en la docencia universitaria y el Espacio Europeo de Educación Superior: cuestiones clave para su puesta en práctica. *Revista de Educación*. <https://redined.educacion.gob.es/xmlui/handle/11162/72311>
- Pérez Ramírez, C. (1993). Evolución histórica de la Educación física. *Apunts*, 33, 24–38.
- Pérez Samaniego, V. (2021). *Tradiciones cualitativas en la investigación sobre Educación Física*. In Monforte, J., & Pérez Samaniego, V. (Coords.), *Investigación cualitativa en educación física, actividad física y deporte* (pp. 29–44). INDE.
- Pérez-Samaniego, V., Fuentes Miguel, J., & Devís Devís, J. (2011). El análisis narrativo en la educación física y el deporte. *Movimiento*, 17(4), 11–42. <https://doi.org/10.22456/1982-8918.24402>

- Perrenoud, P. (1998). *La evaluación: una herramienta de aprendizaje*. In P. Perrenoud, *La evaluación de los alumnos: de la producción de la excelencia a la regulación de los aprendizajes* (pp. 15–30). Ediciones Colihue.
- Phillips, S. R., Marttinen, R., & Mercier, K. (2017). Fitness assessment: Recommendations for an enjoyable student experience. *Strategies*, 30(5), 19-24. <https://doi.org/10.1080/08924562.2017.1344168>
- Piedra, J., García-Pérez, R., Latorre, Á., & Quiñones, C. (2013). Género y educación física. Análisis de buenas prácticas coeducativas. Profesorado. *Revista de Currículum y Formación de Profesorado*, 17(1), 221-241. <https://www.redalyc.org/articulo.oa?id=56726350014>
- Prieto, A. (2015). Los paradigmas de evaluación en educación física evaluation paradigms in physical education. Multiárea. *Revista de didáctica*, Núm. 7 (2015): 110-130. <https://doi.org/10.18239/mard.v0i7.665>
- Prieto-Castillo, D., Guevara-Toledo, C., & Célleri-Gomezcoello, Á. (2023). Más allá de la defensa y el ataque: La evaluación como alternativa de aprendizaje. *Universidad del Azuay*. <https://doi.org/10.33324/ceazuay.178>
- Poblete Gálvez, C. A., & Moreno Doña, A. (2023). La mirada del género femenino en la educación física. Génesis de una historia en Chile. *Géneros*, 22(17), 151–166. <https://revistasacademicas.ucol.mx/index.php/generos/article/view/1160>
- Pole, K. (2009). Diseño de metodologías mixtas. Una revisión de las estrategias para combinar metodologías cuantitativas y cualitativas. *Renglones. Revista arbitrada en ciencias sociales y humanidades*, 60, 37-42. <https://rei.iteso.mx/items/ecb843fa-1c4b-4255-a633-f877b11a22bb>
- Popenici, S.A.D., Kerr, S. Exploring the impact of artificial intelligence on teaching and learning in higher education. *RPTEL* 12, 22 (2017). <https://doi.org/10.1186/s41039-017-0062-8>
- Portes-Júnior, M., Novoa-Ramírez, C.A., Pérez-Gutiérrez, P. A., Roco-Espinoza, G. A., & Valdés-Retamal, S.E. (2022). Contenidos curriculares, aprendizajes esperados e instrumentos de evaluación utilizados por los docentes de Educación Física en Enseñanza Media de la comuna

- de Talca: un estudio piloto. *Journal of Movement & Health*, 19(2) 1–9.
[https://doi.org/10.5027/jmh-Vol19-Issue2\(2022\)art150](https://doi.org/10.5027/jmh-Vol19-Issue2(2022)art150)
- Quintero Hernández, J. (2022). *El recreo como recurso en la Educación Física bajo perspectiva de género* (Trabajo de fin de grado, Universidad de La Laguna). Repositorio institucional ULL.
<http://riull.ull.es/xmlui/handle/915/27662>
- Quirós, G. (2021). Gordofobia: existencia de un cuerpo negado. Análisis de las implicaciones subjetivas del cuerpo gordo en la sociedad moderna. *Revista Latinoamericana de Derechos Humanos*, 32(1), 1–10. <https://doi.org/10.15359/rldh.32-1.7>
- Ramírez-Quintero, F. A., & Rojas-Gámez, C. A. (2025). Investigación, contexto y aplicaciones de la evaluación educativa: Una mirada del desarrollo científico en Latinoamérica. *Revista de Investigación, Administración e Ingeniería*, 13(1), 143–149.
<https://doi.org/10.15649/2346030X.4857>
- Rannau, J. (2020). Educación física en Chile: Hacia una transdisciplinariedad desde el currículum y la colaboración pedagógica. *Praxis educativa*, 24(2), 1–17.
<https://doi.org/10.19137/praxiseducativa-2020-240210>
- Ravela, P. (2009). Consignas, devoluciones y calificaciones: los problemas de la evaluación en las aulas de educación primaria en América Latina. *Páginas De Educación*, 2(1), 49–89.
<https://doi.org/10.22235/pe.v2i1.703>
- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez Rizo, F., Aylwin, M., & Wolff, L. (2016). Las Evaluaciones Educativas que América Latina Necesita. *Revista Iberoamericana De Evaluación Educativa*, 1(1). <https://doi.org/10.15366/riee2008.1.1.004>
- Real Academia Española. (2014). *Evaluar*. En *Diccionario de la lengua española* (23.^a ed.).
<https://dle.rae.es/evaluar>
- Reay, D., & Wiliam, D. (1999). ‘I’ll be a nothing’: Structure, agency and the construction of identity through assessment. *British Educational Research Journal*, 25(3), 343–354.
<https://doi.org/10.1080/0141192990250305>

- Roca Marín, D., & Pardo Quiles, V. (2023). Análisis de la idoneidad del microrrelato en la divulgación científica. *Círculo de Lingüística Aplicada a la Comunicación*, 95, 287–299. <https://doi.org/10.5209/clac.74307>
- Red de Evaluación Formativa y Compartida en Educación (REFyCE). (s.f.). *Presentación*. <https://redevaluacionformativa.wordpress.com/>
- Rodríguez, C., Romero Garrido, D., & Espinosa-Valenzuela, D. (2025). ¿Mayor aprendizaje o inflación de calificaciones?: Un análisis comparado de los sistemas evaluativos internos y externos en la enseñanza secundaria. *Revista Enfoques Educativos*, 22(1), 74–96. <https://doi.org/10.5354/2735-7279.2025.74683>
- Rodríguez, I., del Valle, S., & De la Vega, R. (2018). Revisión nacional e internacional de las competencias profesionales de los docentes de Educación Física. *Retos*, 34, 383-388. <https://doi.org/10.47197/retos.v0i34.58609>
- Rodriguez, F., Monteiro, D., Teixeira, D. S., & Cid, L. (2020). The relationship between teachers and peers' motivational climates, needs satisfaction, and physical education grades: An AGT and SDT approach. *International Journal of Environmental Research and Public Health*, 17(17), 6145. <https://doi.org/10.3390/ijerph17176145>
- Rodríguez, F. R., Gatica, C. C., Bernal, J. L., & Almagià, A. A. F. (2017). Evaluación de la educación física escolar en Enseñanza Secundaria. *Retos: nuevas tendencias en educación física, deporte y recreación*, (31), 76-81. <https://doi.org/10.47197/retos.v0i31.49097>
- Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (Eds.). (2011). e-Evaluación orientada al e-Aprendizaje estratégico en Educación Superior. *Narcea*. <https://www.researchgate.net/publication/364489709>
- Rodríguez-Ortiz, J. K. (2003). La evaluación del desempeño: aproximación desde una perspectiva de género. *Ciencia y Sociedad*, 28(2), 187-205. <https://www.redalyc.org/articulo.oa?id=87028201>
- Rodríguez Sabiote, C., Lorenzo Quiles, O., & Herrera Torres, L. (2005). Teoría y práctica del análisis de datos cualitativos. Proceso general y criterios de calidad. *Revista Internacional de*

- Ciencias Sociales y Humanidades, SOCIOTAM*, XV(2), 133-154.
<http://biblioteca.udgvirtual.udg.mx/jspui/handle/123456789/1038>
- Román San Martín, M. S. (2023). *Mejoras a partir de la implementación de evaluación formativa y compartida en oralidad: Implementación de estrategia evaluativa con foco en evaluación formativa y compartida para la mejora de habilidades orales discursivas* (Tesis de maestría, Pontificia Universidad Católica de Chile). Repositorio UC.
<https://doi.org/10.7764/tesisUC/EDU/75056>
- Rosales, M. M. (2014). *Proceso evaluativo: evaluación sumativa, evaluación formativa y Assesment su impacto en la educación actual*. Congreso Iberoamericano de Ciencia, Tecnología, Innovación y Educación, (Vol. 4, p. 662) 1–13.
<http://www.oei.es/congreso2014/memoriactei/662.pdf>
- Rubin, H. J., & Rubin, I. S. (2020). *Qualitative interviewing: The art of hearing data* (4th ed.). SAGE Publications.
- Ryan, R. M., & Deci, E. L. (2019). *Brick by brick: The origins, development, and future of self-determination theory*. In A. J. Elliot (Ed.), *Advances in Motivation Science* (Vol. 6, pp. 111-156). Cambridge, MA: Elsevier Inc. <https://doi.org/10.1016/bs.adms.2019.01.001>
- Sabirón, F., & Arraiz, A. (2016). Aprendiendo de la evaluación: decálogo para la evaluación auténtica de competencias profesionales a través del Portafolio. *Revista Iberoamericana de Evaluación Educativa*, 6(1). <https://doi.org/10.15366/rie2013.6.1.008>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144. <https://doi.org/10.1007/BF00117714>
- Şahin, M. D. y Ozturk, G. (2019). Mixed method research: Theoretical foundations, designs and its use in educational research. *International Journal of Contemporary Educational Research*, 6(2), 301–310. <https://doi.org/10.33200/ijcer.574002>
- Saiz-González, P., Iglesias, D., Coto-Lousas, J., & Fernández-Río, J. (2025). Exploring pre-service teachers' negative memories of Physical Education. *Espiral. Cuadernos del Profesorado*, 18(37), 47–59. <https://doi.org/10.25115/ecp.v18i37.10033>

- Sánchez Carracedo, F., Soler, A., Martín, C., López, D., Ageno, A., Cabré, J., Garcia, J., Aranda, J., & Gibert, K. (2018). Competency maps: An effective model to integrate professional competencies across a STEM curriculum. *Journal of Science Education and Technology*, 27(5), 448–468. <https://doi.org/10.1007/s10956-018-9735-3>
- Sánchez, G. Q. (2021). Gordofobia: existencia de un cuerpo negado. Análisis de las implicaciones subjetivas del cuerpo gordo en la sociedad moderna. *Revista Latinoamericana de Derechos Humanos*, 32(1). <https://doi.org/10.15359/rldh.32-1.7>
- Sánchez Alvarado, A., & Orozco Alvarado, J. (2020). Las competencias y la evaluación. Hacia un modelo de evaluación auténtica de los aprendizajes. *Revista Científica de FAREM-Esteli*. <https://doi.org/10.5377/farem.v0i32.9225>
- Sánchez, G. Q. (2021). Gordofobia: existencia de un cuerpo negado. Análisis de las implicaciones subjetivas del cuerpo gordo en la sociedad moderna. *Revista Latinoamericana de Derechos*. <https://doi.org/10.15359/rldh.32-1.7>
- Sánchez-Hernández, N., Soler-Prat, S. y Martos-García, D. (2022). La Educación Física desde dentro. El discurso del rendimiento, el currículum oculto y las discriminaciones de género, *Ágora para la Educación Física y el Deporte*, 24, 46-71 <https://doi.org/10.24197/aefd.24.2022.46-71>
- Sánchez Martín, C., García López, L., & Reyes, L. (2022). Beneficios de un enfoque activista de la educación física en la participación deportiva de niñas de Educación Primaria. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 46, 946–954. <https://doi.org/10.47197/retos.v46.94629>
- Sánchez, M., & Vega, J. C. (2003). Algunos aspectos teórico-conceptuales sobre el análisis documental y el análisis de información. *Ciencias de la Información*, 34(2), 49–57.
- Sánchez-Oliva, D., Viladrich, C., Amado, D., González Ponce, I., y García-Calvo, T. (2014). Predicción de los comportamientos positivos en educación física: una perspectiva desde la Teoría de la Autodeterminación. *Revista de Psicodidáctica*, 19(2), 387–406. <https://doi.org/10.1387/RevPsicodidact.7911>

- Sánchez Alvarado, A. F., & Orozco Alvarado, J. C. (2020). Las competencias y la evaluación. Hacia un modelo de evaluación auténtica de los aprendizajes. *Revista Científica Estelí*, (32), 3–14. <https://doi.org/10.5377/farem.v0i32.9225>
- Sandoval Rubilar, P., Maldonado-Fuentes, A. C., & Tapia-Ladino, M. (2022). Evaluación educativa de los aprendizajes: Conceptualizaciones básicas de un lenguaje profesional para su comprensión. *Páginas De Educación*, 15(1), 49–75. <https://doi.org/10.22235/pe.v15i1.2638>
- Santos Guerra, M. A. (1995). *Un proceso de diálogo, comprensión y mejora*. En *La evaluación: Un proceso de diálogo, comprensión y mejora* (pp. 26–33). Aljibe.
- Santos Guerra, M. Á. (2003). Dime cómo evalúas y te diré qué tipo de profesional y de persona eres. *Enfoques Educativos*, 5(1), 69–80.
- Santos Guerra, M. A. (2014). *La evaluación como aprendizaje*. Narcea.
- Santos, J. M., Ortiz, E. & Marín, S. (2018). Índices de variación de la nota debidos a la evaluación continua. Contrastación empírica en la enseñanza universitaria. *Cultura y Educación*, 30(3), 491-527. <https://doi.org/10.6018/rie.329781>
- Sarni, M. y Corbo, J.L. (2020). La evaluación en Educación Física: más allá del control social. *Educación Física y Deporte*, 39(1), 137-165. <https://doi.org/10.17533/udea.efyd.v39n1a07>
- Saugy, J. J., Drouet, O., Millet, G. P., & Lentillon-Kaestner, V. (2020). A systematic review on self-determination theory in physical education. *Translational Sports Medicine*, 3(2), 134-147. <https://doi.org/10.1002/tsm2.121>
- Scharagrodsky, P. A. (2004). Juntos pero no revueltos: la educación física mixta en clave de género. *Cadernos de Pesquisa*, 34, 59-76. https://www.memoria.fahce.unlp.edu.ar/art_revistas/pr.14080/pr.14080.pdf
- Scriven, M. (1967). *The methodology of evaluation*. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Rand McNally.

- Secchi, J. D., García, G. C., & Arcuri, C. R. (2016). ¿Evaluar la condición física en la escuela?: Conceptos y discusiones planteadas en el ámbito de la educación física y la ciencia. *Enfoques*, 28(1), 67–92. <https://www.redalyc.org/articulo.oa?id=25955333004>
- Sérgio, M. (2006). Motricidade humana: o itinerário de um conceito. *Revista da Sociedade de Pesquisa Qualitativa em Motricidade Humana*, 6(1), 15-25. <https://doi.org/10.29181/2594-6463-2022-v6-n1-p15-25>
- Serra, P., Prat, M., da Silva, A., Soler, S., & Silva, A. M. (2020). Género y currículum de formación del profesorado en Educación Física: un diálogo entre Brasil y España. *Revista Iberoamericana de Educación*, 82(2), 175–194. <https://doi.org/10.35362/rie8223637>
- Sevil-Serrano, J., Aibar, A., Abós, Á., & Generelo, E. (2020). Improving motivation for physical activity and physical education through a school-based intervention. *Journal of Experimental Education*, 90, 383-403. <https://doi.org/10.1080/00220973.2020.1764466>
- Silva Batatina, M. de la P. (2017). El transitar en la investigación narrativa y su empleo en la construcción de teoría. *Revista de Investigación*, 41(91), 124–142. <http://www.redalyc.org/articulo.oa?id=376156277008>
- Silva Castillo, R., Pérez Pacheco, L., & Rodríguez Villamizar, J. (2024). *Evaluación sistémica en Educación Física. Lecturas: Educación Física y Deportes*, 29(318), 1–10. <https://doi.org/10.46642/efd.v29i318.7670>
- Silva, R. J. R., Cunha, F. M. P., Da Silva, A. M. C., & Batista, P. M. F. (2024). Empowering students in physical fitness through assessment for learning in physical education. *Journal of Physical Education and Sport*, 24(5), 1174-1182. <https://doi.org/10.7752/jpes.2024.05135>
- Slingerland, M., Weeldenburg, G., & Borghouts, L. (2024). Formative assessment in physical education: teachers' experiences when designing and implementing formative assessment activities. *European Physical Education Review*, 0(0). <https://doi.org/10.1177/1356336X241237398>

- Smith, B. (2017). Generalizability in qualitative research: misunderstandings, opportunities and recommendations for the sport and exercise sciences. *Qualitative Research in Sport, Exercise and Health*, 10(1), 137–149. <https://doi.org/10.1080/2159676X.2017.1393221>
- Smith, B., McGannon, K. R., & Williams, T. L. (2015). Ethnographic creative nonfiction: Exploring the whats, whys and hows. In *Ethnographies in sport and exercise research* (pp. 59-73). Routledge. <https://durham-repository.worktribe.com/output/1631084>
- Smith, L., & Pérez, J. (2023). Insights of High School Students in Learning Folk Dance during Physical Education. *International Journal of School Health*, 10(2), 155-167. <https://doi.org/10.51386/25815946/ijsms-v6i3p107>
- Sonllewa, M., Martínez, S., & Monjas, R. (2018). Los Procesos de Evaluación y sus Consecuencias. Análisis de la Experiencia del Profesorado de Educación Física. *Estudios Pedagógicos XLIV*, N°, 2, 329–351. <https://doi.org/10.4067/S0718-07052018000200329>
- Starck, J., Richards, K. A. R., & O’Neil, K. (2018). A conceptual framework for assessment literacy: Opportunities for physical education teacher education. *Quest*, 70(4), 519–535. <https://doi.org/10.1080/00336297.2018.1465830>
- Stieg, R. y dos Santos, W. (2021). Concepciones de evaluación y Educación Física en la formación del profesorado en Argentina, Chile, México y Uruguay. *Calidad en la Educación*, 55, 82–119. <https://doi.org/10.31619/caledu.n55.1052>
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. Jossey-Bass.
- Su, J., Pu, X., Yadav, K., & Subramnaiyan, M. (2022). A physical education teacher motivation from the self-evaluation framework. *Computers & Electrical Engineering*, 98, 107772. <https://doi.org/10.1016/j.compeleceng.2022.107772>
- Suazo González, P. (2023). *La evaluación formativa y compartida como estrategia para promover el aprendizaje autorregulado en estudiantes con talento académico* [Tesis de maestría, Pontificia Universidad Católica de Chile]. <https://doi.org/10.7764/tesisuc/edu/75093>

- Sundaresan, N., Dashoush, N., & Shangraw, R. (2017). Now that we're 'well rounded,' let's commit to quality physical education assessment. *Journal of Physical Education, Recreation & Dance*, 88(8), 35–38. <https://doi.org/10.1080/07303084.2017.1356770>
- Svennberg, L., Meckbach, J., Redelius, K. (2018). Swedish PE teachers struggle with assessment in a criterion-referenced grading system. *Sport, Education and Society*, 23(4), 381–393. <https://doi.org/10.1080/13573322.2016.1200025>
- Talanquer, V. (2015). La importancia de la evaulacion formativa. *Educacion Quimica*, 26(3), 177–179. <https://doi.org/10.1016/j.eq.2015.05.001>
- Talero, E., & Guarnizo, N. (2022). Las prácticas evaluativas de los docentes de educación física en Villavicencio. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, 44, 775–783. <https://doi.org/10.47197/retos.v44i0.91995>
- Taras, M. (2005). Assessment, summative and formative, some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466–478. <https://doi.org/10.1111/j.1467-8527.2005.00307.x>
- Tejada, J., & Ruiz, C. (2015). Evaluación de competencias profesionales en educación superior. *Educación XXI*, 19(1), 17–38. <https://doi.org/10.5944/educxx1.12175>
- Tendinha, R., Alves, M. D., Freitas, T., Appleton, G., Gonçalves, L., Ihle, A., Gouveia, É. R., & Marques, A. (2021). Impact of Sports Education Model in Physical Education on students' motivation: A systematic review. *Children*, 8(7), 588. <https://doi.org/10.3390/children8070588>
- Thorburn, M., & MacDonald, D. (2011). Standards, criteria and instrumentalism: The influence of high-stakes assessments on physical education curricula. *Physical Education and Sport Pedagogy*, 16(2), 227–241. <https://doi.org/10.3390/su13020578>
- Tinning, R. (2017). *Physical education, curriculum and pedagogy*. Routledge.
- Tirado-Olivares, S., González-Calero, J. A., Cózar-Gutiérrez, R., & Toledano, R. (2021). Gamificando la Evaluación: Una Alternativa a la Evaluación Tradicional en Educación Primaria. *REICE*, 19(4). <https://doi.org/10.15366/reice2021.19.4.008>

- Toro-Suaza, C. A., & Chaverra, B. E. (2022). ¿Cómo evalúan los profesores de Educación Física? Análisis de su acción evaluativa. *RECIE. Revista Caribeña De Investigación Educativa*, 6(1), 28–43. <https://doi.org/10.32541/recie.2022.v6i1.pp28-43>
- Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education*, 38(3), 323–342. <https://doi.org/10.1080/03054985.2012.689693>
- Torrebadella-Flix, X. (2017). La historia de la educación física escolar en España. Una revisión bibliográfica transversal para incitar a una historia social y crítica de la educación física. *Espacio, Tiempo y Educación*, 4(1), 1–41. <https://www.redalyc.org/articulo.oa?id=477455339014>
- Towner, M. C., Keath, A. M., & Wyant, J. D. (2025). Citations and significance: Unraveling the impact of assessment research in physical education through a bibliometric lens. *Quest*, 77(2), 245–262. <https://doi.org/10.1080/00336297.2025.2527170>
- Townsley, M., & McNamara, S. (2021). “I thought I was supposed to get an A in PE!” Successes and challenges of teachers and administrators implementing standards-based grading in physical education. *Studies in Educational Evaluation*, 70, 101012. <https://doi.org/10.1016/j.stueduc.2021.101012>
- Trainor, L. R., & Bundon, A. (2021). Developing the craft: Reflexive accounts of doing reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 13(5), 705–726. <https://doi.org/10.1080/2159676X.2020.1840423>
- Trejo González, H. (2024). Evaluación auténtica en contexto universitario a través de portafolios electrónicos de aprendizaje. *Revista Portuguesa de Educação*, 37. <https://doi.org/10.21814/rpe.29644>
- Trejo González, H. (2024). Instrumento de evaluación para el desarrollo de cursos en entornos virtuales de enseñanza-aprendizaje. *EDUTEC. Revista Electrónica de Tecnología Educativa*, (79), 30–45. <https://doi.org/10.21556/edutec.2022.79.2353>

- Troncoso, C. E., & Daniele, E. G. (2013). Las entrevistas semiestructuradas como instrumentos de recolección de datos: una aplicación en el campo de las ciencias naturales. *Anuario Digital de Investigación Educativa*, (14). <https://revistas.bibdigital.uccor.edu.ar/index.php/adiv/article/view/3313>
- Trussi, L. M. (2020). Estereotipos y representaciones sociales en las clases de Educación Física: una perspectiva de género. *Lecturas: Educación Física Y Deportes*, 25(263), 148-159. <https://doi.org/10.46642/efd.v25i263.1910>
- Tyler, R. W. (2009). *Basic principles of curriculum and instruction*. En D. J. Flinders & S. J. Thornton (Eds.), *The curriculum studies reader* (pp. 69–77). Routledge.
- UNESCO. (2005). *Declaración Universal sobre Bioética y Derechos Humanos*. https://unesdoc.unesco.org/ark:/48223/pf0000146180_spa
- Valencia Peris, A., & Lizandra Mora, J. (2018). Cambios en la representación social de la educación física en la formación inicial del profesorado. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 34, 230-235. <https://doi.org/10.47197/retos.v0i34.60144>
- Valdivia Yábar, S. V., & Fernández Guillén, M. E. del C. (2020). La evaluación formativa en un contexto de renovación pedagógica: Prácticas al servicio del éxito. *Actualidades Investigativas En Educación*, 20(1), 26. <https://doi.org/10.15517/aie.v20i1.40159>
- Vallejo Ruiz, M., & Molina Saorín, J. (2014). La evaluación auténtica de los procesos educativos. *Revista Iberoamericana de Educación*, 64. <https://doi.org/10.35362/rie640403>
- Véliz, B., Ketterer, B., & Thuillier, C. (2020). Technical-instrumental rationality in physical education: A qualitative study in Chilean context. *Revista Espacios*, 41(4), Artículo 19. <https://www.revistaespacios.com/a20v41n04/20410419.html>
- Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 43(5), 840–854. <https://doi.org/10.1080/02602938.2017.1412396>

- Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogical imaginaries and the limits of educational technology. *Learning, Media and Technology*, 45(2), 105–112. <https://doi.org/10.1080/17439884.2020.1761641>
- Zabalza, M. A. (2000). *La enseñanza universitaria: el escenario y sus protagonistas*. Madrid: Narcea.
- Zabalza, M. A., & Lodeiro, L. (2019). El desafío de evaluar por competencias en la universidad. *Revista Iberoamericana de Evaluación Educativa*, 12(2), 29–48. <https://doi.org/10.15366/rie2019.12.2.002>
- Zacarias, I. G. (2018). Las políticas de evaluación educativa en América Latina. *Revista Fuentes*, 20(2), 29–35. <https://doi.org/10.12795/revistafuentes.2018.v20.i2.02>
- Zakaria, Z., Mohamed, A.M., Abdul Ghani, D. Z. y Ali, M.A. (2024). Assessment under duress: Challenges to physical education assessments and the emergence of compromised norms. *International Journal of Academic Research in Progressive Education and Development*, 13(3), 756–774. <https://doi.org/10.6007/IJARPED/v13-i3/21675>
- Zapata-Ros, M. (2024). IA generativa y ChatGPT en Educación: Un reto para la evaluación y ¿una nueva pedagogía? *Revista Paraguaya de Educación a Distancia (REPED)*, 5(1), 12–44. <https://doi.org/10.56152/reped2024-vol5num1-art2>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhu, X. (2015). Student perspectives of grading in physical education. *European Physical Education Review*, 21(4), 409–420. <https://doi.org/10.1177/1356336X15569628>
- Zubillaga-Olague, M., Cañadas, L., & Moura, A. (2023). Assessment aspects of student's motor skills and assessment tools in physical education. *Apunts. Educación Física y Deportes*, 153, 27–38. [https://doi.org/10.5672/apunts.2014-0983.es.\(2023/3\).153.03](https://doi.org/10.5672/apunts.2014-0983.es.(2023/3).153.03)



VNIVERSITAT DE VALÈNCIA

ANEXOS

Anexo 1. Acta comité de ética



Comité de Ética

NºCEC 38-21

ACTA DE EVALUACIÓN PROYECTO DE INVESTIGACIÓN

En Temuco, a 13 de diciembre de 2021, el Comité Ético Científico de la Universidad Autónoma de Chile representado por la Dra. Marlén Barreto, informan han sesionado en revisión expedita, el Protocolo de Investigación, según se indica:

Título del Proyecto	Evaluación en la asignatura de Educación física en Chile. Percepción del alumnado y profesorado.
Investigador Responsable	Fernando Muñoz Sepúlveda
Institución	Programa de Doctorado en Didácticas Específicas de la Universidad de Valencia-España. Académico Universidad Autónoma de Chile.
Decisión	APROBADO
Fecha de Presentación al Comité de Ética Institucional	24 de noviembre de 2021
Documentos Revisados por Comité de Ética	• Formulario de Solicitud. • Proyecto in extenso • Carta de compromiso con Declaración de No conflicto de interés • Consentimientos Informados • Carta de apoyo institucional • CVs investigadores • Instrumentos de medición

El Comité Ético Científico de la Universidad Autónoma de Chile certifica, que después de evaluar el estudio presentado por el académico de esta casa de estudios y candidato a Doctor en Didácticas Específicas por la Universidad de Valencia-España, Sr.

Fernando Muñoz Sepúlveda, cumple con las siguientes consideraciones éticas:

Consideraciones Generales:

El presente estudio es de tipo cualitativo, enfocado en comprender los procesos de evaluación en la asignatura de Educación Física y Salud en Chile desde la mirada del alumnado y del profesorado. Para esto, se caracterizarán las experiencias vividas sobre la evaluación de las asignaturas de Educación Física y Salud en Chile en el periodo escolar, en estudiantes de pregrado de diferentes carreras de la Universidad Autónoma de Chile sede Talca, indagando en la influencia de estas experiencias en la motivación y percepciones por parte de los estudiantes. Además, se pretende complementar la información aportada por los estudiantes con las experiencias y propuestas de profesores que imparten dichas asignaturas.

Ante estas consideraciones se estima que el proyecto cumple con:

1. El estudio no vulnera la dignidad de los participantes involucrados, asegura su derecho a anonimato y garantiza la protección de la confidencialidad de los datos.
2. El estudio se enmarca en los principios de respeto a los Derechos Humanos garantizando en los procedimientos, metodologías y procesos investigativos.
3. El protocolo de investigación se ajusta a los estándares éticos y es coherente con los preceptos contenidos en la Declaración Universal sobre Bioética y Derechos Humanos (UNESCO, 2005), así como a la declaración de Helsinki, y el Proyecto se ajustó a la normativa ética propia de la disciplina en estudio.

En consideración, el Comité Ético Científico de la Universidad Autónoma de Chile APRUEBA la certificación del proyecto de investigación titulado **“Evaluación en la asignatura de Educación física en Chile. Percepción y Adecuación desde el aula”**, cuyo investigador responsable es el candidato a Doctor **Fernando Muñoz Sepúlveda**.

Para constancia firma,


Dra. Marté Barreto



Presidenta Comité Ético Científico Institucional

El Comité Ético Científico de la Universidad Autónoma de Chile está conformado por 11 miembros representados en sus tres sedes

Anexo 2. Consentimiento informado estudiantado



CONSENTIMIENTO INFORMADO

El propósito de este documento es informarle y solicitar su participación voluntaria en el proyecto de investigación titulado “*Evaluación en la asignatura de Educación Física y Salud en Chile: percepciones del profesorado y del alumnado.*” dirigido por el Prof. Dr© Fernando Muñoz Sepúlveda (fernando.munoz@uautonoma.cl), en el marco de la tesis doctoral del programa de la Universidad de Valencia en Didácticas específicas.

Como estudiante de la facultad de Educación de la Universidad Autónoma de Chile, sede Talca, he sido invitado a participar en este proyecto enfocado a Analizar y comprender los procesos de evaluación en la asignatura de Educación Física y Salud en Chile, desde la mirada del profesorado y del alumnado. Para la consecución de los objetivos se llevará a cabo un estudio cualitativo de enfoque Narrativo, en las carreras de Pedagogía de la facultad de Educación de la Universidad Autónoma de Chile, sede Talca, en donde se recolectará información para dar respuesta a los objetivos.

Con los resultados de este estudio se pretende robustecer la línea de investigación de la carrera en

“Didáctica de la Educación física”, teniendo estrecha relación con el quehacer docente en el aula. En base a los resultados, se espera comprender el fenómeno evaluativo de la disciplina a nivel escolar, aportando valiosa información científica, incluso a nivel de asignaturas de la formación de los futuros profesores de Educación física, como, por ejemplo; Bases para la Didáctica y Evaluación del aprendizaje, Didáctica y Evaluación para la motricidad I y II.

Pedimos su colaboración para realizar las siguientes actividades:

- Redactar de forma honesta, sus experiencias vividas en el periodo escolar, sobre las temáticas propuestas de la asignatura de Educación Física y Salud.
- Participar, si es convocado, a participar de grupos focales para complementar sus relatos narrativos. Esta instancia será grabada por el investigador principal Fernando Muñoz Sepúlveda con grabadora de voz, sin imagen.

El tiempo necesario para realizar las evaluaciones será de 45-60 minutos, en la primera etapa de carácter Narrativo. En una instancia diferente, se coordinará una fecha para participar de una reunión que tendrá por duración de 45 a 60 minutos. Por tanto, el tiempo total de dedicación a participar del estudio será de 120 minutos totales aproximadamente. Se

contemplarán los siguientes protocolos sanitarios; Sanitización de manos al ingresar o comenzar la recolección de datos, uso permanente de mascarilla, distanciamiento social y ventilación en espacios cerrados. Además, se adecuarán las instancias de recolección de datos a formato virtual o presencial, según lo permitan las restricciones sanitarias que entregue el MINSAL, respecto a aforos.

Los datos y grabaciones se almacenarán durante 4 años, tiempo en el que se puede extender el programa de doctorado.

Con el fin de proteger la identidad de los participantes, los resultados serán codificadas y sólo el investigador conocerá mi nombre. La información obtenida será de uso científico y no será divulgada con otros fines. Esa información podrá ser publicada en revistas científicas y se mantendrá mi identidad estrictamente confidencial. Los resultados generales serán expuestos en congresos nacionales e internacionales, sin nunca revelar la identidad de los participantes.

No hay costos para mí y tampoco beneficios económicos a los participantes o calificación asociada a alguna asignatura de su formación.

Mi participación es completamente voluntaria. Si retiro mi consentimiento para participar, la información ya obtenida no será usada.

Declaro que he recibido una copia de este documento llamado “consentimiento informado” para leerlo antes de firmar y que se me han explicado los detalles del protocolo.

Don/Doña _____, con RUT número _____, comprendiendo las características del proyecto de investigación expuesto y habiendo podido preguntar todo aquello que he considerado oportuno, ACEPTO LIBREMENTE COLABORAR CON EL ESTUDIO MENCIONADO ANTERIORMENTE.

Y para que así conste, firmo el presente consentimiento.

Fecha:

Firma:

Si necesita contactar con el equipo de investigación puede hacerlo a través del e-mail fernando.munoz@uautonoma.cl o al teléfono (71) 2735581.

Si tuviese dudas respecto a la conducta de investigación, contactarse con la secretaria ejecutiva del comité de Ética de la Universidad, al número: (56-45) 2895249.



Proyecto código CEC 38-21 aprobado por el CEC UA con fecha 13 de diciembre 2021
comite.etica@uautonoma.cl

Anexo 3. Consentimiento informado profesorado



CONSENTIMIENTO INFORMADO

El propósito de este documento es informarle y solicitar su participación voluntaria en el proyecto de investigación titulado “*Evaluación en la asignatura de Educación Física y Salud en Chile: percepciones del profesorado y del alumnado.*” dirigido por el Prof. Dr© Fernando Muñoz Sepúlveda (fernando.munoz@uautonoma.cl), en el marco de su tesis doctoral del programa de la Universidad de Valencia en Didácticas específicas.

Como Profesor de la asignatura Educación Física y Salud en su establecimiento educacional y con al menos 3 años ejerciendo, ha sido invitado a participar en este proyecto enfocado a analizar y comprender los procesos de evaluación en la asignatura de Educación Física y Salud en Chile, desde la mirada del profesorado y del alumnado. Para la consecución de los objetivos se llevará a cabo un estudio cualitativo de enfoque fenomenológico, observando clases y luego en base a esos datos, desarrollar entrevistas semiestructuradas.

Con los resultados de este estudio se pretende robustecer la línea de investigación de la carrera en

“Didáctica de la Educación física”, teniendo estrecha relación con el quehacer docente en el aula. En base a los resultados, se espera comprender el fenómeno evaluativo de la disciplina a nivel escolar, aportando valiosa información científica, incluso a nivel de asignaturas de la formación de los futuros profesores de Educación física, como, por ejemplo; Bases para la Didáctica y Evaluación del aprendizaje, Didáctica y Evaluación para la motricidad I y II.

Pedimos su colaboración para realizar las siguientes actividades:

- Estar dispuesto a ser observado durante un semestre en sus clases de Educación Física. Esta instancia no será grabada, ya que se utilizará un diario de campo para recolectar la información necesaria.
- Responder de forma verbal a una entrevista semiestructurada. Esta instancia será grabada por el investigador principal Fernando Muñoz Sepúlveda con grabadora de voz, sin imagen

El tiempo requerido en las entrevistas será entre 60 a 90 minutos. Respecto a la observación de clases, la duración dependerá de la clase, en al menos 3 instancias durante el semestre. Se contemplarán los siguientes protocolos sanitarios; Sanitización de manos al ingresar o comenzar la recolección de datos, uso permanente de mascarilla, distanciamiento social y ventilación en espacios cerrados. Además, se adecuarán las instancias de recolección de

datos a formato virtual o presencial, según lo permitan las restricciones sanitarias que entregue el MINSAL, respecto a aforos.

Con el fin de proteger la identidad de los participantes, los resultados serán codificadas y sólo el investigador conocerá mi nombre. La información obtenida será de uso científico y no será divulgada con otros fines. Esa información podrá ser publicada en revistas científicas y se mantendrá mi identidad estrictamente confidencial. Los resultados generales serán expuestos en congresos nacionales e internacionales, sin nunca revelar la identidad de los participantes.

No hay costos para mí y tampoco beneficios económicos a los participantes.

Mi participación es completamente voluntaria. Si retiro mi consentimiento para participar, la información ya obtenida no será usada.

Declaro que he recibido una copia de este documento llamado “consentimiento informado” para leerlo antes de firmar y que se me han explicado los detalles del protocolo.

Don/Doña _____, con RUT número _____, comprendiendo las características del proyecto de investigación expuesto y habiendo podido preguntar todo aquello que he considerado oportuno, ACEPTO LIBREMENTE COLABORAR CON EL ESTUDIO MENCIONADO ANTERIORMENTE.

Y para que así conste, firmo el presente consentimiento.

Fecha:

Firma:

Si necesita contactar con el equipo de investigación puede hacerlo a través del e-mail fernando.munoz@uautonoma.cl o al teléfono (71) 2735581.

Si tuviese dudas respecto a la conducta de investigación, contactarse con la secretaria ejecutiva del comité de Ética de la Universidad, al número: (56-45) 2895249.



Proyecto código CEC 38-21 aprobado por el CEC UA con fecha 13 de diciembre de 2021

comite.etica@uautonoma.cl

Anexo 4. Instrumento narrativo

Instrumento Narrativo sobre experiencias y percepciones del alumnado sobre la Evaluación en la asignatura de Educación Física y Salud en Chile.

Desde el programa de Doctorado en Didácticas específicas de la Universidad de Valencia, estamos haciendo un estudio con el fin de analizar y comprender los procesos de evaluación en la asignatura de Educación Física y Salud en Chile, desde la percepción de estudiantes que hayan pasado por la asignatura y desde su relato narrativo en base a sus vivencias en edad escolar.

Si necesita contactar con el equipo de investigación puede hacerlo a través del e-mail fernando.munoz@uautonoma.cl o al teléfono (71) 2735581.

Si tuviese dudas respecto a la conducta de investigación, contactarse con la secretaria ejecutiva del comité de Ética de la Universidad, al número: (56-45) 2895249.

1. Correo *

2. He leído y firmado el consentimiento informado

Marca solo un óvalo.

☐ Si

☐ No

Caracterización

3. Correo electrónico

4. Carrera

Marca solo un óvalo.

- ☐ Pedagogía en Educación Física
- ☐ Pedagogía en Educación Básica
- ☐ Pedagogía en Lengua Castellana y Comunicación
- ☐ Pedagogía en Educación Diferencial
- ☐ Pedagogía en Inglés
- ☐ Pedagogía en Educación Parvularia
- ☐ Pedagogía en Historia, Geografía y Ciencias Sociales

5. Edad

6. Género

Marca solo un óvalo.

- ☐ Femenino
- ☐ Masculino
- ☐ Prefiero no
- ☐ decirlo

Otro:

7. Podría hacer una breve descripción de los establecimientos de educación básica y media por los que estudió. Nombre del establecimiento, ubicación, cambios, etc. (Ejemplo; Colegio San Rafael de la ciudad de Los Angeles en enseñanza básica, enseñanza media en el Liceo Abate Molina de la ciudad de Concepción).

NARRATIVA

En esta etapa es necesario que te remontes a tu edad escolar, recordando vivencias desde primero básico a cuarto medio, con respecto a la EVALUACIÓN en la asignatura de Educación Física o Educación Física y Salud. La idea es que nos relates, en base a tus experiencias reales, los procesos de Evaluación que vivenciaste en la asignatura. El estilo de relato depende de tu comodidad, puedes escribir un relato general o momentos específicos que nos ayudarán a construir la realidad que experimentaste. Es muy relevante, además, describir cuales fueron tus percepciones sobre esas instancias de Evaluación en la asignatura, qué sentimientos fueron los que generaron esas experiencias, ya sea desde el agente evaluador, instrumento, clima, sentido, etc.

8. Una guía podría ser responder a las preguntas generales sobre la EVALUACIÓN EN LA ASIGNATURA DE EDUCACIÓN FÍSICA; ¿Cuándo?, ¿Qué?, ¿Quién?, ¿Cómo?, ¿Por qué? Y sobre todo...¿Cómo te hizo sentir ? Escribe en primera persona, entre más completa sea la narrativa, es mejor para la investigación, no hay un mínimo de palabras, pero trata de abordar todos los elementos que para ti son relevantes.

Anexo 5. Guion de grupo focal para estudiantado

Focus Group sobre experiencias y percepciones del alumnado sobre la Evaluación en la asignatura de Educación Física y Salud en Chile

El moderador deberá iniciar la reunión con el siguiente relato

Buenos/as días/tardes y muchas gracias por participar de esta conversación que está enmarcada en la segunda etapa de la investigación para el programa de doctorado en didácticas específicas de la Universidad de Valencia sobre la Evaluación en la asignatura de Educación Física y Salud en Chile. Mi nombre es Fernando Muñoz Sepúlveda y seré el moderador de este grupo de discusión. Les haré algunas preguntas y esperamos recibir sus respuestas de la forma más completa y honesta posible. Javiera Verdugo tomará notas de esta reunión y también utilizará una grabadora, para luego poder completar su registro y reflejar lo más fielmente posible las opiniones vertidas. Les aseguramos que esta grabación no se difundirá en ninguna instancia y será utilizada sólo con fines de transcripción. Sus identidades se mantendrán en confidencialidad, solo registraremos opiniones y no personas.

Temáticas a abordar:

1.- Clase de Educación física en básica y media

1.1- La metodología

- ¿Qué metodologías vivenciaron frecuentemente en sus clases de EF en edad escolar? Relacionemos esas metodologías con la motivación

- ¿Qué opinan de los procesos metodológicos vivenciados en la clase de EF?

1.2- La planificación

- ¿Las clases vivenciadas en edad escolar, ¿qué se transmite en base a la planificación, la secuencia didáctica, la conexión de una clase con otra?

1.3- Deserción de la clase

- En los relatos narrativos, se ha mencionado el deseo de abandonar la clase, ¿qué relación tiene la metodología con el deseo de deserción?

2.- Evaluación en Educación Física

2.1- El sentido de la Evaluación

- ¿Qué sentido perciben ustedes desde la Evaluación en la asignatura de Educación Física?

2.2- Los instrumentos

- ¿Qué instrumentos se utilizaban con mayor frecuencia? ¿Ustedes conocían los instrumentos, criterios o indicadores? (el instrumento fue socializado, compartido por el profesorado ...)

2.3- Retroalimentación

¿Qué importancia, prevalencia, relevancia tenía la retroalimentación en las clases de EF?

2.4- Emociones en la evaluación

- ¿Las emociones provocadas en las instancias de evaluación en EF, han tenido alguna repercusión en su actual afinidad por la actividad física? ¿Cómo es esa relación?

- En algunos relatos ha surgido el miedo. Ese miedo especialmente por la Gimnasia Artística. ¿Podemos profundizar aquello? O si surgen otros momentos, deportes, etc.

2.6- Evaluación formativa

- ¿Qué experiencias de evaluación ustedes podrían destacar, en donde se favorezca el aprendizaje? Más adecuada. ¿Hubo aprendizaje en su experiencia de EF?

3.- Discriminación en Educación física

3.1- Diferencias de género cual es la relevancia del género en las clases de EF.

- Profundicemos en las diferencias de género percibidas en la clase de EF en edad escolar.

3.2- “Gordofobia” y discriminación por rendimiento

- Ha surgido el término “gordofobia”, o que solo se veían favorecidos los que tenían una mejor condición física. ¿Pueden profundizar al respecto?

Anexo 6. Guion de entrevista semiestructurada para el profesorado

El moderador deberá iniciar la reunión con el siguiente relato

Buenos/as días/tardes y muchas gracias por participar de esta entrevista que está enmarcada en la tercera etapa de la investigación para el programa de Doctorado en Didácticas Específicas de la Universidad de Valencia, España, sobre la Evaluación en la asignatura de Educación Física y Salud en Chile. Mi nombre es FERNANDO MUÑOZ SEPÚLVEDA y seré el entrevistador. Les haré algunas preguntas y esperamos recibir sus respuestas de la forma más completa y honesta posible. También se utilizará una grabadora, para luego poder completar su registro y reflejar lo más fielmente posible las opiniones vertidas. Les aseguro que esta grabación no se difundirá en ninguna instancia y será utilizada sólo con fines de transcripción. Dicha transcripción solo será usada con los fines académicos relacionados con el proyecto de tesis. Una vez transcrita, les será enviada para que comprueben que se recogen fielmente sus respuestas. Además, pueden negarse a contestar cualquiera de las preguntas que se viertan. Su identidad se mantendrá en confidencialidad, solo registraremos opiniones y no personas como se menciona en el consentimiento informado que han accedido a firmar.

Objetivo de la entrevista:

Caracterizar los procedimientos de evaluación y calificación que el profesorado aplica en la actualidad, identificando el sentido, dimensiones y tipos de evaluación.

Foco	Preguntas
Caracterización de personas participantes y preguntas de familiarización	- Consentimiento informado-Characterización en Google forms: https://forms.gle/Gx3iSH5cvujUegJBA - ¿Cómo ha sido su experiencia como profesor(a) de EF desde la vuelta a la presencialidad? - ¿Cómo definirías la evaluación? - ¿Para qué sirve? - ¿Qué opina de las propuestas de Evaluación del curriculum nacional?
Descripción del proceso evaluación y calificación	- ¿Cómo planificas tu evaluación? ¿En qué te basas, en la programación anterior, en algún libro, etc.? - ¿Cómo evalúa/califica en Educación Física? ¿Utiliza algún instrumento? ¿Lo conoce el alumnado? En qué momentos, ¿cómo es el proceso de creación de dichos instrumentos? ¿Participa el alumnado en su evaluación? ¿La calificación está asociada a la evaluación? - ¿Cuántas calificaciones le exigen por semestre en EF?, ¿Podría tomar un nivel de ejemplo y describir de qué se trataron el semestre pasado? - Para este año, ¿el plan de evaluación será el mismo que el del año pasado para ese nivel?, podría mencionar las diferencias o similitudes.

	- Nos puede comentar los resultados, promedios generales. ¿Qué pasa si son bajos? ¿Qué espera el centro educativo o la familia de los resultados de calificaciones en la asignatura? - ¿Genera situaciones de evaluaciones que no conlleven una calificación? ¿De qué se tratan?
Identificar Tipos de evaluación	- ¿Propone al inicio de un proceso algún tipo de evaluación para conocer el nivel de conocimientos del estudiantado? En caso afirmativo, ¿para qué le sirve? - ¿Cómo participa el estudiantado en los procesos de evaluación?
Aplicación de Test de condición física	- ¿Los aplica? ¿Establece alguna diferencia por género? - ¿Desde qué nivel/curso? - Podría comentar qué prueba son los que utiliza y para qué. - ¿Tienen vinculación con la calificación?, ¿cómo es ese proceso? ¿Según tablas? Etc. - ¿Cuáles son las reacciones del alumnado cuando presenta los test de condición física? ¿Cómo los viven los alumnos y las alumnas? ¿Y aquéllos con menor competencia motriz o discapacidad?
Evaluación según Dimensión	- ¿Cómo aborda las dimensiones conceptuales, procedimentales y actitudinales en sus procesos evaluativos? ¿Siempre están presentes en cualquier tipo de contenido las tres dimensiones?
Evaluación por contenido curricular	- ¿Podría comentar cómo evalúa la Unidad de Habilidades Motrices Básicas? (Si es profesor(a) de primer ciclo. - ¿Podría comentar cómo evalúa la Unidad de deportes o juegos predeportivos? (Si es profesor(a) de segundo ciclo en adelante. - ¿Podría comentar cómo evalúa la Unidad de Danzas folklóricas? (Si es profesor(a) de segundo ciclo en adelante.
Influencia de experiencias de evaluación	¿Cuáles son las semejanzas y diferencias entre las formas de evaluar que usted propone y las que usted vivificó siendo escolar?
Inclusión	- ¿Cómo ha podido dar respuesta a la diversidad de estudiantes en sus procesos de evaluación? Desde Necesidad Educativas Especiales, ¿obesidad, discapacidad motora, nivel de competencia motriz, género etc.? ¿Ha encontrado dificultades o limitaciones al respecto?
Libertad	- Podría dar algunas recomendaciones al profesorado que enseña “evaluación” a los futuros docentes de EF. Lo que desde su experiencia ha sido relevante. - ¿Desea aportar cualquier tipo de comentario, observación o crítica respecto a los procesos evaluativos que usted desee?

Se le agradece su participación y el tiempo que, desinteresadamente, nos ha regalado. su valioso tiempo y Le recordamos que le enviaremos la transcripción de esta entrevista por si desea realizar alguna apreciación al respecto.

Anexo 7. Registro oficial de calificaciones nacionales desde MINEDUC



Esquema de registro bases de Rendimiento escolar 2007, 2011-2022, notas de los estudiantes en cada subsector de su plan de estudios.

Base pública con MRUN

Unidad responsable:	Unidad de Estadísticas Centro de Estudios Ministerio de educación
Fuente de la información:	Año 2007: Registro de Estudiantes de Chile (RECH) Años 2011-2022: Sistema de Información General de Estudiantes (SIGE)
Formato de la base de datos:	.CSV con codificación UTF-8 ¹

Número de observaciones por año

Año escolar	Número de registros	Número de Establecimientos ²
2022	32.793.724	8.903
2021	32.410.770	8.964
2020	25.196.613	9.040
2019	32.227.180	9.068
2018	30.612.736	9.160
2017	30.188.037	9.243
2016	30.236.779	9.312
2015	30.002.869	9.416
2014	29.971.878	9.518
2013	29.976.971	9.612
2012	28.266.796	9.733
2011	28.744.899	9.792
2007	30.811.114	10.025

¹ Hasta el año 2017 la codificación de la base de datos es de entorno local.

² El número de establecimientos de este archivo puede no coincidir con el número de establecimientos de la base de datos de Rendimiento Escolar, ya que en este se consideran sólo los establecimientos con al menos un alumno con nota válida en algún subsector de su plan de estudios.