
ACREDITACIÓN Y DESEMPEÑO EN EVALUACIONES ESTANDARIZADAS DE LA FORMACIÓN INICIAL DOCENTE EN EDUCACIÓN FÍSICA CHILENA

FRANCISCO GALLARDO-FUENTES, BASTIAN CARTER-THUILLIER,
JORGE GALLARDO-FUENTES, CAROLINA MARTÍNEZ-ANGULO
Y SEBASTIÁN PEÑA-TRONCOSO

RESUMEN

La formación inicial docente en Educación Física en Chile se encuentra fuertemente regulada por sistemas de aseguramiento de la calidad, donde la acreditación constituye un indicador clave del desempeño institucional. En este contexto, el objetivo del estudio fue analizar la relación entre los años de acreditación de las carreras de Pedagogía en Educación Física y el desempeño de sus estudiantes en la Evaluación Nacional Diagnóstica (END), considerando los estándares disciplinares y las dimensiones temáticas evaluadas. Se utilizó un diseño cuantitativo, transversal y comparativo, basado en el análisis de los resultados de 1.085 estudiantes de último año de formación inicial docente que rindieron la END 2024. Los

resultados muestran un comportamiento no lineal en función de los años de acreditación, destacando que las instituciones con cinco años presentan los mayores niveles de logro en la mayoría de los estándares. En contraste, las instituciones con seis años no evidencian un rendimiento consistentemente superior, mientras que aquellas con tres y cuatro años presentan desempeños más bajos y homogéneos. En conclusión, los hallazgos sugieren que la acreditación no se asocia de manera directa con mejores resultados de aprendizaje, lo que evidencia la necesidad de complementar estos sistemas con indicadores centrados en el desempeño estudiantil y en la calidad efectiva de la formación docente.

Introducción

Los sistemas de educación superior en diversos contextos han incorporado mecanismos de aseguramiento de la calidad como respuesta a la masificación, la

diversificación institucional y las demandas de rendición de cuentas (Harvey y Williams, 2010). En este contexto, la acreditación se ha consolidado como uno de los principales dispositivos de regulación externa de las instituciones, orientado a movilizar los procesos de autorregulación y a promover, mediante

estándares mínimos de calidad, la instalación de procesos de mejora continua (Eaton, 2012).

Desde una perspectiva conceptual, la acreditación puede entenderse como un proceso sistemático de evaluación externa que valida el cumplimiento de criterios previamente definidos,

PALABRAS CLAVE / Acreditación / Aseguramiento de la Calidad / Educación Física / Formación Inicial Docente /

Recibido: 03/04/2026. Modificado: 29/04/2026. Aceptado: 04/05/2026.

Francisco Gallardo-Fuentes (Autor de correspondencia). Doctor en Investigación Transdisciplinar en Educación. Magíster en Educación, Licenciado en Educación. Profesor de Educación Física. Docente, Universidad de Los Lagos, Chile y Universidad Andrés Bello, Chile. Dirección: Universidad de Los Lagos. Av. Fuchslocher 1305, Osorno. Región de Los Lagos, Chile. e-mail: fgallardo@ulagos.cl, f.gallardofuentes@uandresbello.eu. ORCID: 0000-0002-3985-157X.

Bastían Carter-Thuillier. Doctor en Investigación Transdisciplinar en Educación, Máster en Ciencias Sociales para la Investigación en Educación, Licenciado en Educación, Profesor de Educación Física. Académico. Departamento de Educación, Universidad de Los Lagos, Osorno, Chile; Facultad de Educación y Humanidades, Universidad Andrés Bello, Santiago, Chile. Dirección: Universidad de Los Lagos. Av. Fuchslocher 1305, Osorno. Región de Los Lagos, Chile. e-mail: bastiancarter@gmail.com. ORCID: 0000-0001-7342-4179.

Jorge Gallardo-Fuentes. Magíster en Recursos Naturales. Profesor, Escuela de Medicina Veterinaria, Facultad de Medicina y Ciencias de la Salud, Universidad Mayor, Temuco. e-mail: jorgegallardo-fuentes@gmail.com. ORCID: 0000-0002-1701-173X.

Carolina Martínez-Angulo. Magíster en Educación, mención Licenciada en Educación, Universidad de los Lagos, Chile. Directora docente, carrera Técnico Deportivo Universitario, Universidad de los Lagos. e-mail: carolina.martinez@ulagos.cl.

Sebastián Peña-Troncoso. Profesor de Educación Física. Máster en Innovación, Evaluación y Calidad de Educación Física y Doctor en Ciencias de la Educación. Académico, Instituto de Ciencias de la Educación, Universidad Austral de Chile y Profesor invitado. Universidad SEK. e-mail: sebastian.pena@uach.cl.

los cuales abarcan dimensiones como gestión institucional, docencia, resultados de aprendizaje e investigación (Stensaker y Harvey, 2010). En este marco, en la dimensión vinculada al aprendizaje, los resultados de las pruebas estandarizadas pueden constituir un aporte a la mejora de los procesos de enseñanza-aprendizaje (Encarnación, 2022). Sin embargo, la literatura ha advertido que estos sistemas no están exentos de tensiones, particularmente en lo relativo a su capacidad de real para reflejar la calidad educativa de las instituciones (Newton, 2010).

En el caso chileno, el aseguramiento de la calidad se institucionaliza formalmente con la creación de la Comisión Nacional de Acreditación (CNA), organismo encargado de evaluar y acreditar tanto instituciones como programas (Geoffroy Pitta, 2014). La Ley N.º 20.129 y sus modificaciones posteriores establecen que la acreditación tiene como propósito certificar la calidad de los procesos internos y los resultados formativos de las instituciones de educación superior.

En este contexto, en Chile resulta particularmente relevante el carácter obligatorio de la acreditación para las carreras de pedagogía, lo que refleja una preocupación del Estado por asegurar estándares mínimos en la Formación Inicial Docente (FID) (Fernández-Cofré *et al.*, 2021). En este marco, los años de acreditación (que pueden oscilar entre 0 y 7 años) se han transformado en un indicador sintético de calidad, utilizado tanto para la regulación como para la toma de decisiones por parte de estudiantes y políticas públicas.

No obstante, diversos estudios cuestionan la capacidad de los años de acreditación para capturar de manera precisa la complejidad de la calidad educativa, señalando que estos pueden reflejar, más bien, condiciones estructurales y organizativas que resultados efectivos de aprendizaje (Andreani *et al.*, 2020; Bernasconi, 2015). En Chile, de forma paralela al desarrollo de los sistemas de acreditación, se han implementado evaluaciones estandarizadas orientadas a medir resultados de aprendizaje en la FID, como la END, la cual se ha posicionado como uno de los principales instrumentos en esta línea, evaluando competencias disciplinares y pedagógicas de estudiantes próximos a egresar (Maturana-Castillo *et al.*, 2025).

La END se estructura en estándares que buscan operacionalizar los conocimientos y habilidades esperadas en el profesorado en formación, lo que permite generar información comparable a nivel nacional (Brouckaert *et al.*, 2025). En el caso de Educación Física, estos

estándares se organizan en dos grandes dimensiones: fundamentos teóricos y capacidades de implementación pedagógica, tal como se observa en el apartado de metodología del estudio (Flores-Ferro *et al.*, 2025).

Desde la literatura internacional, las evaluaciones estandarizadas han sido ampliamente utilizadas como mecanismos de *accountability*, permitiendo la comparación entre instituciones y la orientación de políticas educativas (Loeb y Byun, 2022). Sin embargo, también han sido objeto de críticas por su potencial reduccionismo y por no capturar dimensiones complejas del aprendizaje (Au, 2007).

En teoría, sistemas de acreditación robustos deberían asociarse con mejores resultados en evaluaciones estandarizadas, dado que promueven mejoras en los procesos formativos (Abduvalievich, 2024). Sin embargo, la evidencia disponible muestra resultados mixtos, ya que algunos estudios sugieren una relación positiva entre acreditación y desempeño estudiantil (Singh-Ahluwalia, 2023), mientras que otros indican que dicha relación es débil o inconsistente (Harvey, 2005). Esto se debe, en parte, a que la acreditación evalúa procesos institucionales, mientras que las pruebas estandarizadas miden resultados individuales, lo que implica niveles de análisis distintos (Salinas, 2024). En el contexto chileno, esta tensión se vuelve particularmente relevante, ya que los años de acreditación son utilizados como garantía de calidad, sin que exista necesariamente evidencia concluyente de que se traduzcan en mejores resultados en evaluaciones como la END (Koch *et al.*, 2025).

El análisis de los resultados de evaluaciones estandarizadas en relación con la acreditación requiere considerar diversos elementos; entre ellos, reconocer que dichos resultados reflejan no solo la calidad institucional, sino también factores individuales, contextuales y curriculares (Trow, 2007).

En este sentido, en el ámbito de la FID, emerge la necesidad de avanzar hacia modelos de evaluación que integren de manera más coherente los procesos formativos, la evaluación auténtica y los resultados estandarizados, en línea con enfoques de evaluación formativa y desarrollo profesional docente (Darling-Hammond *et al.*, 2017). Si bien la acreditación constituye un mecanismo clave de aseguramiento de la calidad, su relación con los resultados de aprendizaje no es directa ni unívoca (Findlay, 2024).

A pesar del desarrollo de estudios sobre acreditación y aseguramiento de la calidad en educación superior, así como de investigaciones que han

mostrado relaciones débiles o no lineales entre acreditación y resultados de aprendizaje, persiste un vacío específico en el campo de la FIPEF en Chile. En particular, existe escasa evidencia empírica que examine, de manera integrada, cómo los años de acreditación de las carreras se relacionan con el desempeño de sus estudiantes en la END, considerando el análisis desagregado por estándares disciplinares y dimensiones temáticas. En este sentido, el presente estudio busca contribuir no solo a corroborar la existencia de una relación no lineal, sino también a precisar cómo dicha relación se expresa en un ámbito disciplinar específico, aportando evidencia empírica relevante para discutir el alcance real de la acreditación como indicador de calidad en la FID.

Por todo lo anterior, el objetivo del presente estudio fue analizar la relación entre los años de acreditación de las carreras de Pedagogía en Educación Física en Chile y el desempeño de sus estudiantes en la Evaluación Nacional Diagnóstica.

Metodología

Desde el punto de vista metodológico, el estudio se enmarca en un enfoque cuantitativo, de corte transversal y con un diseño comparativo (Hernández *et al.*, 2006). Las variables analizadas corresponden a los resultados en los dos temas evaluados y en los seis estándares disciplinares de la END de la FID chilena en Educación Física, expresados como proporciones de logro:

Tema 1: Sustentos teóricos de la disciplina:

Estándar 1: Comprende los referentes teóricos que apoyan su práctica pedagógica disciplinar.

Estándar 2: Conoce las manifestaciones de la motricidad como medio de enseñanza.

Estándar 3: Comprende la importancia del ejercicio físico en el desarrollo de hábitos de una vida activa saludable.

Tema 2: Programación e implementación:

Estándar 4: Conoce las características motrices, cognitivas, afectivas y sociales del estudiante en situación de acción motriz.

Estándar 5: Organiza la didáctica de la Educación Física.

Estándar 6: Conoce el proceso evaluativo propio de la disciplina.

Adicionalmente, se incorporó una variable externa correspondiente a los años de acreditación institucional, la cual es obligatoria para todas las carreras de pedagogía impartidas en la educación superior chilena. Los datos

sobre acreditación se obtuvieron a partir de una base de datos secundaria, la cual fue depurada para excluir instituciones sin acreditación vigente.

El tratamiento de los años de acreditación como tramos discretos (3, 4, 5 y 6 años) respondió a una decisión analítica orientada a comparar categorías efectivamente presentes en la base y con suficiente viabilidad estadística para su contraste. Esta forma de agrupación permitió estandarizar el análisis entre niveles formalmente diferenciados de acreditación, asumiendo que cada tramo representa un umbral institucional distinto dentro del sistema chileno de aseguramiento de la calidad. No obstante, esta decisión también delimita el alcance interpretativo de los resultados, ya que los grupos comparados expresan posiciones administrativas dentro del sistema de acreditación, pero no agotan la heterogeneidad organizacional, curricular o pedagógica existente al interior de cada institución.

La evaluación de la normalidad se realizó mediante la prueba de Shapiro–Wilk, la cual evidenció que los datos no seguían una distribución normal en los distintos grupos. Posteriormente, se realizaron comparaciones entre grupos de acreditación mediante la prueba U de Mann-Whitney para muestras independientes, considerando todas las combinaciones posibles entre los cuatro niveles de acreditación en los que se agruparon las distintas carreras de Pedagogía en Educación Física en Chile (3, 4, 5 y 6 años). Se reportaron los valores del estadístico U, el nivel de significancia (p) y el tamaño del efecto (r), calculado a partir de la transformación del estadístico Z. El tamaño del efecto se interpretó de acuerdo con criterios convencionales, considerando valores cercanos a 0,10 como pequeños, 0,30 como moderados y 0,50 como grandes, lo que permite estimar la relevancia práctica de las diferencias estadísticamente significativas (Cohen, 2013). Todos los análisis fueron realizados en el entorno R (R Core

Team), utilizando principalmente los paquetes rstatix, stats, dplyr y openxlsx para la gestión y exportación de datos.

Resultados

Los resultados analizan los porcentajes de logro de 1.085 estudiantes de último año de la Formación Inicial del Profesorado en Educación Física (FIPEF) chilena que rindieron la END 2024. Al categorizar los datos según institución de dependencia, se observa un patrón diferenciado en el rendimiento en función de los años de acreditación, lo cual se presenta en la Tabla I.

En términos generales, las instituciones con 5 años de acreditación presentan los promedios más altos en la mayoría de los estándares (Est1, Est3, Est4, Est5 y Est6), destacando particularmente en Est1 (M= 0,56), Est5 (M= 0,58) y Est6 (M= 0,59), lo que se asocia con un mayor nivel de desempeño en este grupo. Las instituciones con 3 y 4 años de acreditación muestran rendimientos más bajos y bastante similares entre sí, con valores en Est4 de M= 0,42 y 0,40, respectivamente, configurando el nivel más bajo del conjunto. Por su parte, las instituciones con 6 años de acreditación presentan el valor más alto en Est2 (M= 0,56) y muestran buenos niveles en Est4 (M= 0,47). Estos resultados sugieren la presencia de un patrón parcialmente asociado a los años de acreditación, en el cual los niveles más altos concentran los mejores desempeños, especialmente en el caso de las instituciones con 5 años de acreditación. No obstante, las diferencias observadas son de magnitud moderada, por lo que deben interpretarse con cautela.

En las instituciones con 3 años de acreditación (Tabla I), los promedios se distribuyen de manera relativamente homogénea entre los estándares, con valores que oscilan entre M= 0,42 (Estándar 4) y M= 0,57 (Estándar 6). Destacan los mayores desempeños

en los Estándares 5 y 6, mientras que el Estándar 4 presenta el valor más bajo, lo que sugiere mayores dificultades en este dominio. Para las instituciones con 4 años de acreditación, se observa un patrón similar, con promedios entre M= 0,40 (Estándar 4) y M= 0,55 (Estándar 5). En este grupo, los resultados tienden a ser levemente inferiores en comparación con otros niveles, manteniéndose una menor performance en el Estándar 4 y valores intermedios en el resto de los estándares.

En el caso de las instituciones con 5 años de acreditación, se evidencian los mejores desempeños generales, con promedios que van desde M= 0,45 (Estándar 4) hasta M= 0,59 (Estándar 6). Este grupo presenta los valores más altos en la mayoría de los estándares, particularmente en los Estándares 1, 5 y 6, lo que sugiere un rendimiento más consolidado en estos ámbitos. Por su parte, las instituciones con 6 años de acreditación muestran un comportamiento diferenciado, con promedios entre M= 0,47 (Estándar 4) y M= 0,56 (Estándares 2 y 5). En este grupo destaca el mayor desempeño en el Estándar 2, mientras que en otros estándares los resultados son comparables o levemente inferiores a los observados en el grupo de 5 años.

En la Tabla II se presentan las comparaciones entre años de acreditación para cada estándar.

En términos generales, los resultados evidencian que las diferencias entre grupos según años de acreditación no son homogéneas entre estándares (Tabla II) y que, en la mayoría de los casos, los tamaños del efecto son de magnitud baja a moderada. En el Estándar 1, se identifican diferencias estadísticamente significativas en las comparaciones entre 3 y 5 años, así como entre 4 y 6 años y entre 5 y 6 años ($p \leq 0,01$). En estos casos, los tamaños del efecto oscilan entre $r = 0,13$ y $r = 0,17$, lo que corresponde a magnitudes pequeñas a cercanas a moderadas,

TABLA I
RENDIMIENTO EN ESTÁNDARES POR AÑO DE ACREDITACIÓN

Años de acreditación	n	Estándar 1		Estándar 2		Estándar 3		Estándar 4		Estándar 5		Estándar 6	
		M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
3	209	0,49	0,21	0,50	0,20	0,50	0,18	0,42	0,20	0,56	0,21	0,57	0,18
4	430	0,49	0,17	0,48	0,17	0,50	0,15	0,40	0,17	0,55	0,19	0,54	0,18
5	307	0,56	0,18	0,51	0,18	0,52	0,17	0,45	0,18	0,58	0,19	0,59	0,17
6	139	0,50	0,20	0,56	0,18	0,50	0,17	0,47	0,19	0,56	0,19	0,53	0,18

M: media, ds: desviación estándar.

TABLA II
COMPARACIONES POR AÑOS DE ACREDITACIÓN PARA CADA ESTÁNDAR (MANN-WHITNEY)

Comparaciones	Estándar 1			Estándar 2			Estándar 3		
	U	p	r	U	p	r	U	p	r
3 vs 4 años	44491	0,84	0,01	46642	0,43	0,03	44999	0,98	0,00*
3 vs 5 años	26125	0,01*	0,16	30702	0,40	0,04	30044	0,21	0,05
3 vs 6 años	14047	0,60	0,03	11631	0,01*	0,17	14288	0,79	0,01
4 vs 5 años	52912	0,01*	0,17	60247	0,04*	0,08	61069	0,08	0,06
4 vs 6 años	29226	0,69	0,02	22358	0,01*	0,19	29174	0,67	0,02*
5 vs 6 años	24747	0,01*	0,13	17782	0,01*	0,14	22398	0,39	0,04*

Comparaciones	Estándar 4			Estándar 5			Estándar 6		
	U	p	r	U	p	r	U	p	r
3 vs 4 años	47202	0,29	0,04	47097	0,32	0,04	48637	0,09	0,07
3 vs 5 años	29257	0,08	0,08	30493	0,33	0,04	29709	0,15	0,06
3 vs 6 años	12378	0,02*	0,13	14401	0,89	0,01	15861	0,14	0,08
4 vs 5 años	56258	0,01*	0,13	59618	0,02*	0,08	55718	0,01*	0,13
4 vs 6 años	23704	0,01*	0,16	28176	0,31	0,04	30257	0,82	0,01
5 vs 6 años	19913	0,25	0,05	22235	0,47	0,03	25037	0,01*	0,14

U: U de Mann-Whitney, r: tamaño del efecto, p: nivel de significancia estadística, *p: $\leq 0,05$.

sugiriendo diferencias consistentes, pero no de gran magnitud.

Para el Estándar 2, se observa un patrón más marcado de diferencias. Destacan las comparaciones entre 3 y 6 años, 4 y 6 años y 5 y 6 años, todas con significancia estadística ($p \leq 0,04$), y tamaños del efecto que alcanzan valores de hasta $r = 0,19$, lo que indica un efecto cercano a moderado. Esto sugiere que este estándar es particularmente sensible a los años de acreditación institucional.

En el Estándar 3, no se observan diferencias estadísticamente significativas en ninguna de las comparaciones ($p > 0,05$), lo que indica una relativa homogeneidad entre los grupos en este dominio.

En el Estándar 4, se identifican diferencias significativas en las comparaciones entre 3 y 6 años, 4 y 5 años y 4 y 6 años, con valores de $p \leq 0,02$ y tamaños del efecto entre $r = 0,13$ y $r = 0,16$, lo que sugiere diferencias sistemáticas de baja a moderada magnitud. Respecto al Estándar 5, las diferencias son más acotadas, observándose significancia únicamente en la comparación entre 4 y 6 años ($p = 0,02$; $r = 0,08$), lo que indica un efecto pequeño. Finalmente, en el Estándar 6, se evidencian diferencias significativas en las comparaciones entre 4 y 6 años y 5 y 6 años ($p \leq 0,01$), con tamaños del efecto cercanos a $r = 0,13$ – $0,14$, lo que corresponde a magnitudes pequeñas a moderadas.

En cuanto a los temas, las diferencias en el desempeño según los años de acreditación institucional son de magnitud moderada, cuyos resultados se presentan en la Tabla III.

En el Tema 1, se observa una tendencia general al alza a medida que aumentan los años de acreditación, pasando de un promedio de 0,50 en instituciones con 3 años a 0,53 en aquellas con 5 y 6 años. Este patrón sugiere un mejor desempeño en este componente en instituciones con mayor trayectoria acreditada.

En el Tema 2, el comportamiento es menos lineal. Si bien las instituciones con 5 años de acreditación presentan el promedio más alto ($M = 0,54$), las de 6 años muestran un descenso ($M = 0,50$), situándose en niveles similares a las instituciones con 4 años. Esto podría

indicar que el desarrollo de este componente no sigue necesariamente una lógica de acumulación progresiva como en el Tema 1.

Los resultados de las comparaciones múltiples mediante la prueba de Mann-Whitney evidencian diferencias estadísticamente significativas entre varios grupos de años de acreditación, las cuales se presentan en la Tabla IV, aunque con tamaños de efecto generalmente pequeños.

En el Tema 1, se observan diferencias significativas en las comparaciones 3 vs 5 años ($p < 0,001$; $r = 0,13$), 3 vs 6 años ($p = 0,02$; $r = 0,13$), 4 vs 5 años ($p < 0,01$; $r = 0,15$) y 4 vs 6 años ($p < 0,01$; $r = 0,13$). En contraste, no se evidencian diferencias entre 3 vs 4 años ($p = 0,96$) ni entre 5 vs 6 años ($p = 1,00$). Este patrón sugiere que las

TABLA III
RENDIMIENTO EN LOS TEMAS POR AÑO DE ACREDITACIÓN

Años de acreditación	n	Tema 1		Tema 2	
		M	ds	M	ds
3	209	0,50	0,14	0,51	0,16
4	430	0,49	0,11	0,50	0,13
5	307	0,53	0,12	0,54	0,14
6	139	0,53	0,13	0,50	0,15

M: media, ds: desviación estándar.

TABLA IV
COMPARACIONES POR AÑOS DE ACREDITACIÓN Y TEMAS (MANN-WHITNEY)

Comparaciones	Tema 1			Tema 2		
	U	p	r	U	p	r
3 vs 4 años	44826	0,96	0	46421,5	0,5	0,03
3 vs 5 años	27000	0,01*	0,13	27911	0,01*	0,11
3 vs 6 años	12319	0,02*	0,13	14352,5	0,85	0,01
4 vs 5 años	54393	0,01*	0,15	54247	0,01*	0,15
4 vs 6 años	24727,5	0,01*	0,13	28467,5	0,4	0,04
5 vs 6 años	21331	1	0	24044	0,03*	0,1

U: U de Mann-Whitney, r: tamaño del efecto, p: nivel de significancia estadística, *p: $\leq 0,05$.

principales diferencias se concentran entre instituciones con menor acreditación (3 y 4 años) y aquellas con mayor acreditación (5 y 6 años), mientras que entre los grupos superiores el desempeño es equivalente. Los tamaños del efecto, en todos los casos, son pequeños ($r \approx 0,13-0,15$), lo que indica que, si bien las diferencias son estadísticamente significativas, su magnitud es acotada.

En el Tema 2, el patrón es menos consistente. Se identifican diferencias significativas en 3 vs 5 años ($p=0,01$; $r=0,11$), 4 vs 5 años ($p < 0,001$; $r=0,15$) y 5 vs 6 años ($p=0,03$; $r=0,10$). Sin embargo, no se observan diferencias en las comparaciones 3 vs 4 años ($p=0,50$), 3 vs 6 años ($p=0,85$) ni 4 vs 6 años ($p=0,40$). A diferencia del Tema 1, en este caso destaca que el grupo de 5 años presenta diferencias tanto con niveles inferiores como superiores, lo que refuerza la idea de un comportamiento no lineal ya observado en los resultados descriptivos. Nuevamente, los tamaños del efecto son pequeños ($r \approx 0,10-0,15$).

En conjunto, los resultados permiten identificar tres tendencias principales. En primer lugar, la relación entre años de acreditación y desempeño en la END no presenta un comportamiento lineal, ya que las carreras con 5 años de acreditación concentran los resultados más altos y consistentes en la mayor parte de los estándares, sin que las de 6 años muestren una ventaja sistemática.

En segundo lugar, las diferencias más claras tienden a observarse entre los tramos de menor acreditación (3 y 4 años) y los de mayor acreditación (5 y 6 años), aunque con tamaños del efecto predominantemente bajos, lo que sugiere cautela en la interpretación de su alcance práctico.

En tercer lugar, el patrón de resultados no es homogéneo entre dimensiones, puesto que el Tema 1 muestra

una tendencia más estable según años de acreditación, mientras que en el Tema 2 el comportamiento es menos regular. Así, más que evidenciar una asociación directa entre acreditación y rendimiento, los resultados muestran una relación diferenciada según el tipo de competencia evaluada.

Discusión

La relación no lineal entre los años de acreditación y el desempeño en los estándares disciplinares cuestiona la suposición de que mayores niveles de acreditación se traducen automáticamente en mejores resultados de aprendizaje. Si bien las instituciones con 5 años de acreditación concentran los mayores niveles de logro en la mayoría de los estándares, este patrón no se mantiene en las instituciones con 6 años, lo que sugiere que la acreditación podría estar capturando dimensiones organizativas más que efectos directos sobre el aprendizaje. Este hallazgo es consistente con estudios que plantean que los sistemas de aseguramiento de la calidad tienden a evaluar estructuras institucionales y procesos formales, pero no necesariamente garantizan mejoras en los resultados académicos del alumnado que se forma en dichas instituciones (Dill y Beerkens, 2010; Hazelkorn y Mihut, 2021).

Asimismo, la evidencia de desempeños más bajos y homogéneos en instituciones con 3 y 4 años, junto con el hecho de que el grupo de 6 años no supere consistentemente al de 5 años, sugiere la existencia de un efecto de estabilización, fenómeno descrito en estudios recientes sobre calidad en educación superior, donde los incrementos en acreditación no generan mejoras proporcionales en el aprendizaje, sino que más bien reflejan una estandarización de procesos (Akhtar *et al.*, 2025).

Por otra parte, la persistencia de menores resultados en el

Estándar 4 en todos los niveles de acreditación indica la existencia de desafíos estructurales en la formación vinculada a la comprensión del estudiante en contexto pedagógico específico (“Conoce las características motrices, cognitivas, afectivas y sociales del estudiante en situación de acción motriz”). Este tipo de competencias, de carácter más aplicado e integrado, suelen ser menos sensibles a mecanismos de evaluación externa y requieren transformaciones más profundas tanto en las prácticas formativas como en los instrumentos que se utilizan para evaluarlas (Castejón *et al.*, 2018).

Si bien existen diferencias estadísticamente significativas en los porcentajes de logro dependiendo de los años de acreditación, los tamaños del efecto son predominantemente pequeños, lo que sugiere que la acreditación tiene un impacto limitado en términos prácticos sobre el desempeño en los estándares evaluados. Este hallazgo se alinea con investigaciones que sostienen que los sistemas de aseguramiento de la calidad presentan efectos modestos sobre los resultados de aprendizaje, especialmente cuando estos se miden a través de pruebas estandarizadas (Martínez-Zarzuelo *et al.*, 2023).

La mayor sensibilidad del Estándar 2 frente a los años de acreditación sugiere que ciertas áreas del conocimiento disciplinar responden de manera más directa a condiciones institucionales, particularmente aquellas asociadas a contenidos estructurados y evaluables como “conoce las manifestaciones de la motricidad como medio de enseñanza”. En contraste, la ausencia de diferencias en el Estándar 3 (Comprende la importancia del ejercicio físico en el desarrollo de hábitos de una vida activa saludable) indica que algunos aprendizajes presentan una relativa estabilidad, independientemente del nivel de acreditación, lo que ha sido documentado en estudios que

señalan la limitada capacidad de los sistemas de calidad para impactar dimensiones complejas del aprendizaje (Jorre de St Jorre y Oliver, 2018).

No obstante, la baja magnitud de los efectos observados refuerza la necesidad de interpretar estos resultados con cautela, evitando asumir relaciones causales directas entre acreditación y aprendizaje. En este sentido, la evidencia reciente enfatiza que la calidad educativa es un fenómeno multidimensional, donde los resultados dependen de múltiples factores más allá de la acreditación (Benavides y Rodríguez, 2025).

El análisis por temas muestra que la relación entre acreditación y desempeño varía según la naturaleza de las competencias evaluadas, mientras que el tema 1 muestra una tendencia creciente de mejor desempeño asociada a mayores años de acreditación, el tema 2 presenta un comportamiento no lineal, lo que sugiere que las competencias teóricas y prácticas responden de manera diferenciada. Este hallazgo es coherente con la literatura que distingue entre aprendizajes declarativos y competencias prácticas, señalando que estas últimas son menos sensibles a intervenciones estructurales y requieren procesos formativos más complejos (Biesta, 2021; Kump *et al.*, 2015).

El mejor desempeño del grupo de 5 años en ambos temas sugiere la existencia de un punto óptimo en el desarrollo institucional, donde los procesos de mejora aún generan efectos visibles. Sin embargo, el descenso en el grupo de 6 años en el tema 2 refuerza la idea de que la acreditación no garantiza mejoras sostenidas en competencias pedagógicas, lo que ha sido señalado en estudios recientes sobre formación docente, donde se evidencia una desconexión entre políticas de calidad y prácticas pedagógicas efectivas (Menter y Flores, 2021; Shoukat-Ali, 2025).

Los resultados por temas confirman que las diferencias entre niveles de acreditación son estadísticamente significativas, aunque con tamaños de efecto pequeños, lo que sugiere que la acreditación tiene una influencia limitada sobre el desempeño en ambas dimensiones evaluadas. En el caso del tema 1, las diferencias entre niveles bajos y altos de acreditación indican que los procesos de aseguramiento de la calidad podrían estar asociados a mejoras en conocimientos teóricos, lo que coincide con estudios recientes que destacan el impacto de la regulación institucional en dimensiones cognitivas estructuradas (Llanos *et al.*, 2025). Sin embargo, la ausencia de diferencias entre los niveles más altos (5 y 6 años) refuerza la idea

de un efecto de saturación, donde los beneficios de la acreditación tienden a estabilizarse, fenómeno ampliamente documentado en la literatura reciente (Gogoi *et al.*, 2024; Guernon, 2023).

Desde una perspectiva de política pública, estos hallazgos sugieren que el uso de los años de acreditación como indicador sintético de calidad debe ser interpretado con cautela, especialmente cuando se le atribuye un valor orientador para la toma de decisiones institucionales o para la regulación del sistema. El carácter no lineal de los resultados, junto con la baja magnitud de las diferencias observadas, indica que una mayor acreditación no necesariamente se traduce en mejores desempeños estudiantiles en todos los ámbitos evaluados. En este sentido, los sistemas de aseguramiento de la calidad podrían fortalecerse si complementan los indicadores estructurales y de gestión con evidencia más directa sobre procesos formativos, desarrollo de competencias y resultados de aprendizaje, evitando que la acreditación sea leída como un reflejo suficiente de la calidad efectiva de la formación docente.

En el caso del tema 2, la mayor variabilidad observada sugiere que las competencias pedagógicas son menos dependientes de la acreditación y más sensibles a factores contextuales y pedagógicos, lo que resulta consistente con investigaciones en formación docente que enfatizan la importancia de las experiencias prácticas y las interacciones pedagógicas en el desarrollo de competencias profesionales (Amaral *et al.*, 2025; Herrera-Velázquez, 2025).

Conclusiones

A modo de conclusión, los resultados del estudio permiten afirmar que la acreditación, si bien constituye un mecanismo relevante de regulación del sistema de educación superior, no se configura como un indicador de la calidad del aprendizaje en la formación inicial docente en Educación Física. En este sentido, la evidencia muestra que la relación entre años de acreditación y desempeño en la END no es lineal ni consistente, lo que refuerza la necesidad de problematizar su uso en pos de la calidad educativa.

Desde una perspectiva sistémica, los hallazgos sugieren que los sistemas de aseguramiento de la calidad tienden a privilegiar dimensiones más fácilmente medibles, particularmente aquellas asociadas a conocimientos teóricos, en desmedro de competencias más complejas, situadas y contextualizadas. Este fenómeno, ampliamente documentado en la literatura reciente, se ve reflejado en la menor

sensibilidad de ciertos estándares vinculados a la práctica pedagógica y a la comprensión integral del estudiante.

En consecuencia, los resultados refuerzan la necesidad de avanzar hacia modelos de evaluación más integrales en la formación inicial docente, que permitan capturar la complejidad del aprendizaje. Esto implica transitar desde enfoques centrados exclusivamente en el cumplimiento de estándares institucionales hacia sistemas que articulen de manera coherente los procesos formativos, la evaluación auténtica, los resultados de aprendizaje y las prácticas situadas. En conjunto, la evidencia presentada reafirma que, si bien la acreditación cumple un rol relevante en la regulación del sistema, presenta limitaciones como indicador de calidad del aprendizaje, especialmente en dimensiones complejas.

El principal aporte de este estudio radica en ofrecer evidencia empírica situada sobre la relación entre acreditación y desempeño en la FIPEF chilena, mostrando que dicha relación no solo es limitada, sino también diferencial según el tipo de competencia evaluada. Este hallazgo contribuye al campo al matizar interpretaciones lineales sobre la acreditación y al poner en discusión su alcance como proxy de calidad formativa. A partir de ello, se proyecta la necesidad de desarrollar investigaciones que incorporen variables curriculares, pedagógicas e institucionales complementarias, así como estudios comparativos y longitudinales que permitan comprender con mayor profundidad los mecanismos mediante los cuales los sistemas de aseguramiento de la calidad inciden (o no) en los resultados de aprendizaje.

A partir de los hallazgos del estudio, se propone que futuras investigaciones avancen hacia enfoques más complejos y multivariados que permitan comprender de manera más integral la relación entre aseguramiento de la calidad y aprendizaje, incorporando variables que complementen el rendimiento en las evaluaciones estandarizadas como la END. En particular, se sugiere incorporar diseños longitudinales que permitan analizar trayectorias formativas, así como integrar variables institucionales, pedagógicas y contextuales que influyen en el desempeño estudiantil. Asimismo, resulta clave profundizar en el análisis de competencias prácticas y situadas, utilizando metodologías mixtas que complementen los datos estandarizados con evidencia cualitativa.

AGRADECIMIENTOS

El presente artículo está vinculado al proyecto ANID, Programa

FONDECYT Regular, Folio 1241373, titulado: Impacto de la política nacional de estudiantes extranjeros y las orientaciones técnicas para la inclusión educativa de estudiantes extranjeros sobre la formación del profesorado de Educación Física en la macrozona sur de Chile.

REFERENCIAS

- Abduvalievich HG (2024) Quality assurance, audit, and accreditation mechanisms in the most advanced private higher education institutions. *American Journal Of Social Sciences And Humanity Research* 4: 268-273. <https://doi.org/10.37547/ajsshr/Volume04Issue12-24>
- Akhtar S, Bais P, Bahadur PS, Dwivedi M (2025) Worldwide Drifts in Quality Management and Accreditation in Higher Education: En Naim A, Saklani A, Khan SA, Malik PK (Eds.), *Global Perspectives on Quality Management and Accreditation in Higher Education*. IGI Global. pp. 413-430. <https://doi.org/10.4018/979-8-3693-9481-6.ch017>
- Amaral LDM, Costa UBRD, Lima VRD (2025) Capacitação de docentes: Qual é a trajetória recomendada para a estruturação de competências e habilidades na formação dos professores? En Fenner I, Bugarim JP, Monteiro JDR, Oliveira MCD, Galvão SMB, Assis TRDRD (Eds.), *Educação em Debate—Volume 4*. Peleton Science. pp: 168-182. <https://doi.org/10.58976/978-65-986652-1-0.cap12>
- Andreani M, Russo D, Salini S, Turri M (2020) Shadows over accreditation in higher education: Some quantitative evidence. *Higher Education* 79: 691-709. <https://doi.org/10.1007/s10734-019-00432-1>
- Au W (2007) High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher* 36: 258-267.
- Benavides LAM, Rodríguez AVT (2025) Educational Quality: A Systematic Review Using the Prisma Method. *Revista de Gestão Social e Ambiental* 19: e011383. <https://doi.org/10.24857/rgsa.v19n2-100>
- Bernasconi A (2015) *La Educación Superior de Chile: Transformación, Desarrollo y Crisis*. Ediciones UC. Santiago, Chile. 632 pp.
- Biesta G (2021) *World-centred education: A view for the present*. Routledge. Nueva York, EE.UU. 126 pp.
- Brouckaert CL, Cárdenas CB, Garrido PB, Zamora RG (2025) *Manual de uso de resultados de la Evaluación Nacional Diagnóstica de la Formación Inicial Docente*. Ministerio de Educación de Chile, Santiago, Chile. 45 pp.
- Castejón FJ, Santos Pastor ML, Cañadas L (2018) Desarrollo de Competencias Docentes en la Formación Inicial del Profesorado de Educación Física. Relación con los Instrumentos de Evaluación. *Estudios Pedagógicos (Valdivia)* 44: 111-126. <https://doi.org/10.4067/S0718-00752018000200111>
- Cohen J (2013) *Statistical power analysis for the behavioral sciences*. Routledge. 567 pp.
- Darling-Hammond L, Hyster ME, Gardner M (2017) *Effective Teacher Professional Development. Research Brief*. Learning Policy Institute. Washington DC, EEUU. 76 pp.
- Dill DD, Beerkens M (2010) *Public Policy for Academic Quality: Analyses of Innovative Policy Instruments*. Springer. Dordrecht, Países Bajos. 338 pp.
- Eaton J (2012) *An Overview of US Accreditation (revised)*. Council for Higher Education Accreditation. CHEA Washington DC, EE.UU. 12 pp.
- Encarnación AC (2022) The Use of Standardized Tests to Measure Learning: Past, Present, and Future. En Ponce OA, Encarnación ALC, Pagán-Maldonado N, Galán JG, *Introduction to the Philosophy of Educational Research* (1.a ed). River Publishers. Dinamarca. pp. 125-137. <https://doi.org/10.1201/9781003338697-10>.
- Fernández-Cofré MB, Fernández-Quevedo L, Díaz-Navarro M, Jofré-Cáceres P (2021) Respuesta e interpretación a políticas de rendición de cuentas de formación docente en Chile. *Pensamiento Educativo: Revista de Investigación Educativa Latinoamericana* 58 (1). <https://doi.org/10.7764/PEL.58.1.2021.9>.
- Findlay S (2024) A Scoping Review of the Relationships Between Quality Assurance, Accreditation and Program Evaluation in Higher Education. University of Ottawa. ON, Canadá. 44 pp. <https://doi.org/10.20381/RUOR-30229>.
- Flores-Ferro E, Maureira-Cid F, Prat-Lopicich A, Castillo-Retamal F, Bahamondes-Acevedo V, Cortés-Cortés M, Arroyo-Jofré P (2025) Análisis de Resultados de la Evaluación Nacional Diagnóstica en Pedagogía en Educación Física en Chile (2021-2023). *European Public & Social Innovation Review* 11: 1-15. <https://doi.org/10.31637/epsir-2026-2138>.
- Geoffroy Pitta E (2014) Origen y características del sistema de aseguramiento de la calidad de la educación superior chileno. *Revista Electrónica Interuniversitaria de Formación del Profesorado* 17: 49-64. <https://doi.org/10.6018/reifop.17.3.204051>.
- Gogoi AP, Kalita J, Kakati N (2024) Enhancing Care Standards and Accreditation's Impact on Healthcare Quality: En Chaturvedi V, Singh P, Ramachandran A, Aggarwal D (Eds.), *Advances in Healthcare Information Systems and Administration*. IGI Global Scientific Publishing. Hershey, PA, EE.UU. pp. 147-174. <https://doi.org/10.4018/979-8-3693-5523-7.ch005>.
- Guernon E (2023) Reviews of Literature on Accreditation and Quality Assurance. *International Journal of Reviews* 1: 1-57. <https://doi.org/10.5281/ZENODO.8289323>.
- Harvey L (2005) A history and critique of quality evaluation in the UK. *Quality Assurance in Education* 1: 263-276.
- Harvey L, Williams J (2010) Fifteen Years of Quality in Higher Education. *Quality in Higher Education* 16: 3-36. <https://doi.org/10.1080/13538321003679457>.
- Hazelkorn E, Mihut G (2021) *Research Handbook on University Rankings: Theory, Methodology, Influence and Impact*. Edward Elgar Publishing. Cheltenham, Reino Unido. 584 pp.
- Hernández Sampieri R, Fernández C, Baptista P (2006) *Metodología de la Investigación Cuantitativa*. McGraw-Hill. México DF, México. 850 pp.
- Herrera-Velázquez A (2025) Las Competencias Profesionales en los Procesos de Formación Docente Plan 2018 en la Escuela Normal Urbana Cuahtémoc. *Ibero Ciencias - Revista Científica y Académica* 4: 4451-4483. <https://doi.org/10.63371/ic.v4.n3.a351>.
- Jorre de St Jorre T, Oliver B (2018) Want students to engage? Contextualise graduate learning outcomes and assess for employability. *Higher Education Research & Development* 37: 44-57.
- Koch D, López DA, Troncoso E (2025) Las regulaciones de la calidad en la Educación Superior en Chile y su contribución al aseguramiento y mejoramiento de la calidad en la experiencia de una universidad estatal. *Ensaio: Avaliação e Políticas Públicas em Educação* 33: e0254367. <https://doi.org/10.1590/s0104-40362025003304367>.
- Kump B, Moskaliuk J, Cress U, Kimmerle J (2015) Cognitive foundations of organizational learning: Re-introducing the distinction between declarative and non-declarative knowledge. *Frontiers in Psychology* 6: 1489. <https://doi.org/10.3389/fpsyg.2015.01489>.
- Llanos EP, Díaz Flores RR, Aguilar Ventura LJ, Campoverde GAA (2025) The administrative regulation of educational quality in public universities: A study on the regulatory impact on academic processes. *Lex localis - Journal of Local Self-Government* 23: 1328-1340. <https://doi.org/10.52152/801953>.
- Loeb S, Byun E (2022) Educational Accountability and the Role of International Large-Scale Assessments. En Nilsen T, Stancel-Piątak A, Gustafsson J-E (Eds.), *International Handbook of Comparative Large-Scale Studies in Education*. Springer International Handbooks of Education. pp. 75-95. https://doi.org/10.1007/978-3-030-88178-8_6.
- Martínez-Zarzuelo A, Rodríguez-Mantilla JM, García-Domingo B (2023) Evaluación del Impacto de los Sistemas de Acreditación de Educación Superior en el Proceso de Enseñanza-Aprendizaje. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación* 22: 45-64. <https://doi.org/10.15366/reice2024.22.1.003>.
- Maturana-Castillo D, Castillo-Barrientos E, Saba-Henry F, Cisternas-Vejar O, Matus-León, J (2025) Formación del profesorado: La implementación de la Evaluación Nacional Diagnóstica en el sistema chileno: Un balance sobre su uso. *Profesorado, Revista de Currículum y Formación del Profesorado* 29: 1-23. <https://doi.org/10.30827/profesorado.v29i2.30837>.
- Menter I, Flores MA (2021) Connecting research and professionalism in teacher education. *European Journal of Teacher Education* 44: 115-127.
- Newton J (2010) A Tale of Two 'Qualities': Reflections on the Quality Revolution in Higher Education. *Quality in Higher Education* 16: 51-53. <https://doi.org/10.1080/13538321003679499>.
- Salinas L (2024) Análisis de los procesos de evaluación institucional y de acreditación en la Educación Superior de Argentina. *Debate Universitario* 16: 46-60. <https://doi.org/10.59471/debate2024306>.
- Shoukat A (2025) Accreditation, accountability, and advancement: Strategies to elevate teacher training quality in south asia. *Qualitative*

ACCREDITATION AND PERFORMANCE IN STANDARDIZED ASSESSMENTS OF INITIAL TEACHER EDUCATION IN PHYSICAL EDUCATION IN CHILE

Francisco Gallardo-Fuentes, Bastian Carter-Thuillier, Jorge Gallardo-Fuentes, Carolina Martínez-Angulo and Sebastián Peña-Troncoso

SUMMARY

Initial teacher education in Physical Education in Chile is strongly regulated by quality assurance systems, where accreditation constitutes a key indicator of institutional performance. In this context, the aim of this study was to analyze the relationship between the years of accreditation of Physical Education teacher education programs and students' performance in the National Diagnostic Evaluation (END), considering both disciplinary standards and thematic dimensions. A quantitative, cross-sectional, and comparative design was used, based on the analysis of results from 1,085 final-year students who completed

the END 2024. The results show a non-linear pattern according to accreditation years, indicating that institutions with five years achieve the highest levels of performance across most standards. In contrast, institutions with six years do not demonstrate consistently higher results, while those with three and four years show lower and more homogeneous performance. In conclusion, the findings suggest that accreditation is not directly associated with improved learning outcomes, highlighting the need to complement these systems with indicators focused on student performance and the effective quality of teacher education.

ACREDITAÇÃO E DESEMPENHO EM AVALIAÇÕES PADRONIZADAS DA FORMAÇÃO INICIAL DE PROFESSORES EM EDUCAÇÃO FÍSICA NO CHILE

Francisco Gallardo-Fuentes, Bastian Carter-Thuillier, Jorge Gallardo-Fuentes, Carolina Martínez-Angulo e Sebastián Peña-Troncoso

RESUMO

A formação inicial de professores em Educação Física no Chile encontra-se fortemente regulada por sistemas de garantia da qualidade, onde a acreditação constitui um indicador chave do desempenho institucional. Nesse contexto, o objetivo do estudo foi analisar a relação entre os anos de acreditação dos cursos de formação de professores em Educação Física e o desempenho dos estudantes na Avaliação Nacional Diagnóstica (END), considerando os padrões disciplinares e as dimensões temáticas avaliadas. Utilizou-se um delineamento quantitativo, transversal e comparativo, com base na análise dos resultados de 1.085 estudantes do último ano que realizaram a END 2024.

Os resultados mostram um comportamento não linear em função dos anos de acreditação, destacando que as instituições com cinco anos apresentam os maiores níveis de desempenho na maioria dos padrões. Em contraste, as instituições com seis anos não demonstram desempenho consistentemente superior, enquanto aquelas com três e quatro anos apresentam resultados mais baixos e homogêneos. Em conclusão, os achados sugerem que a acreditação não se associa diretamente a melhores resultados de aprendizagem, evidenciando a necessidade de complementar esses sistemas com indicadores centrados no desempenho dos estudantes e na qualidade efetiva da formação docente.

Licencia de uso



Esta obra está bajo una licencia internacional [Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/).