

**INFORME FINAL RESULTADOS COMPARATIVOS POR
CARRERAS Y SUGERENCIAS DE MEJORA EN EL
DESARROLLO DE NUEVAS VERSIONES DE LAS
PRUEBAS DISCIPLINARES**

**PONTIFICIA UNIVERSIDAD CATÓLICA DE
VALPARAÍSO**

Autores:

Pablo Cáceres Serrano

José Miguel Garrido

Verónica Leiva

Livio Núñez

ESCUELA DE PEDAGOGÍA

Noviembre 2017

Resumen

En el siguiente documento se incluyen los resultados finales del proceso de evaluación de estudiantes de primer año en carreras pedagógicas. La información que se dispone corresponde a los resultados de las pruebas disciplinares, creadas por las diferentes unidades académicas en orden a obtener una medición específica del desempeño de sus estudiantes y además, los desempeños transversales o pedagógicos que han sido elaborados desde la Escuela de Pedagogía con el fin de dar cuenta de aquellas capacidades que un estudiante del área debe poseer en el nivel esperado para su condición.

Los resultados denotan que hay diferencias entre los desempeños obtenidos en la prueba de capacidades pedagógicas y la prueba de capacidades disciplinares, siendo más alto el resultado entre carreras como Educación Diferencial y Parvularia, en la primera prueba y mayor entre ciencias básicas, como Física o Química, en la segunda prueba.

La mayoría de los tests cumplen apenas aceptablemente los criterios para ser pruebas fiables o con capacidades de discriminación o dificultad aceptable.

Contenido

Resumen	2
Fundamentos	4
Objetivos	6
Metodología	7
Participantes	7
Diseño	8
Instrumentos	8
Procedimiento	17
Análisis.....	18
Resultados	19
Composición de los instrumentos.	19
Cálculo de la fiabilidad por carrera.	30
Resultados finales por carrera.	31
Síntesis.	41
Bibliografía	42

Fundamentos

A comienzos del año 2017, durante los meses de abril y mayo, se solicitó a la Escuela de Pedagogía, la coordinación en la construcción de instrumentos de medición que permitieran evaluar con qué capacidades o habilidades los estudiantes que ingresan a carreras de carácter pedagógico o que derivaran en título de profesor.

La pretensión fue determinar a través de pruebas aplicadas al inicio de su formación, si los estudiantes tenían las capacidades necesarias para enfrentar los desafíos académicos, sociales y organizacionales a los que se enfrentan cuando ingresan a una carrera de esta índole. La medición fue preparada con el propósito de establecer dos aspectos: a) capacidades: entendidas como aquellas aptitudes que el estudiante debiera poseer al momento de ingresar a la universidad, pueden haber sido formadas durante sus años de colegio o producto de la influencia familiar, por lo que se asume que es capaz de enfrentar ciertos problemas o dificultades de índole académico en un nivel basal y resolverlas; b) actitudes: entendidas como aquella disposición cognitiva, emocional y conductual hacia un objeto u objetos específicos hacia los cuales el individuo adopta una posición y es capaz de favorecer o no ciertos juicios o afirmaciones que están en el orden de dicha posición personal.

Teniendo presente estos dos aspectos, se les solicitó a todas las carreras en esta condición, que construyeran preguntas para evaluar las capacidades que son propias de sus respectivas disciplinas. Esto dio origen a las denominadas pruebas o tests disciplinares, concentrados en la medición de las capacidades que son propias de su ámbito. Por su parte, se consideró también importante evaluar capacidades generales o transversales, que fueron consultadas directamente con todas las carreras para obtener evidencia sólida acerca de la generalidad de dichas aptitudes. Las capacidades transversales dieron origen a la prueba o test pedagógico, que incluyó un apartado destinado a medir las actitudes a favor del trabajo docente y lo que éste implica.

En definitiva, se desarrolló un instrumento compuesto de un test de capacidades pedagógicas transversales a todas las carreras, una escala de actitudes asociadas al mismo propósito y un test específico para cada una de las carreras que participaron en la evaluación, que sólo evaluó capacidades disciplinares.

La evaluación se realizó por vez primera y su pretensión fue explorar las capacidades y actitudes de los estudiantes de primer ciclo. No obstante, el objetivo será posteriormente obtener una medida de aquello que es necesario apoyar, promover o

nivelar en orden a mejorar el desempeño de los estudiantes y en lo posible reducir su deserción.

Se presentó un primer informe en el que se dispusieron los resultados de los análisis efectuados para determinar el funcionamiento de los instrumentos y tomar una decisión respecto de aquellos tests e ítems que requieren una revisión o bien decidir por su eliminación.

En este segundo informe de trabajo, se ofrecen dos tipos de resultados: a) Los ítems que efectivamente han mostrado condiciones psicométricas básicas aceptables para su aplicación, y en consecuencia, la estructura final de la prueba resultante habiendo eliminado los ítems con mal funcionamiento y b) los resultados resúmenes por carrera obtenidos por los estudiantes. Se adjunta además, al final, un anexo con los correspondientes listados de estudiantes por cada carrera y sus desempeños.

Objetivos

Los objetivos de trabajo que guían el trabajo presentado en este informe son los siguientes:

1. Caracterizar el comportamiento psicométrico de los ítems que efectivamente fueron incorporados en las pruebas disciplinares
2. Caracterizar el comportamiento psicométrico de los ítems que efectivamente fueron incorporados en la prueba pedagógica
3. Describir los resultados obtenidos por las diferentes carreras en las pruebas disciplinares que cada unidad académica de la Pontificia Universidad Católica de Valparaíso ha elaborado.
4. Describir los resultados obtenidos por las diferentes carreras en prueba pedagógica elaborada por la Escuela de Pedagogía de la Pontificia Universidad Católica de Valparaíso.

Metodología

Participantes

La población a la que se le aplicó la batería de pruebas fue definida como: Estudiantes que han ingresado durante el año 2017 a las carreras de pedagogía o aquellas que, según curriculum, están en condiciones de obtener el título de educación y/o el grado de licenciado en educación.

El total de evaluados ascendió a 623 estudiantes pertenecientes a 13 carreras. El número de estudiantes por carrera se dispone en la siguiente tabla:

Tabla 1. Número de estudiantes por carrera que rindieron las pruebas

Carrera	Frecuencia	Porcentaje
Biología	31	5.0
Educación Básica	43	6.9
Educación Diferencial	58	9.3
Educación Física	85	13.6
Educación Parvularia	39	6.3
Filosofía	9	1.4
Física	41	6.6
Historia	71	11.4
Inglés	80	12.8
Literatura y Ciencias Lenguaje	63	10.1
Matemáticas	60	9.6
Música	36	5.8
Química	7	1.1
Total	623	100.0

La mayor contribución en frecuencia fue Educación Física, seguido por la carrera de Inglés y posteriormente Historia. Las carreras con menor contribución fueron Química y Filosofía, no necesariamente por haber tenido menor cobertura, sino porque en sí son carreras pequeñas en número de estudiantes.

No hubo un diseño muestral específico pues la pretensión original fue realizar un estudio de la población completa, razón por la cual se evaluó al máximo de estudiantes por disciplina, llegando a una cobertura general aproximada del 90%.

Diseño

El estudio es cuantitativo, transeccional, de tipo instrumental (Montero y León, 2015). Como tal está orientado al análisis del comportamiento de un instrumento en una población específica de sujetos, con objeto de validar las interpretaciones que se obtienen mediante su aplicación a la misma población o poblaciones similares.

La primera etapa de la investigación se correspondió con las actividades propias de un estudio no experimental, realizándose en terreno la recolección de los datos, pues un conjunto de examinadores se presentaron directamente en las aulas con tabletas electrónicas para que los estudiantes respondieran directamente en línea a las pruebas.

Con la información obtenida se procedió a: a) Desarrollar el estudio instrumental, a fin de obtener las propiedades psicométricas de los instrumentos y b) Determinar el desempeño de los estudiantes de las diferentes carreras en función de los ítems seleccionados durante el estudio instrumental. Este segundo informe se centra en los resultados definitivos una vez realizadas las correcciones al instrumento.

Instrumentos

Los instrumentos que se consideran en este documento de análisis han sido elaborados con el fin de medir capacidades, algunas transversales a todas las carreras vinculadas a educación y otras específicas de una disciplina. La prueba de capacidades transversales además está acompañada de una escala gráfica tipo Likert destinada a evaluar las actitudes hacia el trabajo docente.

Atributos generales de construcción.

El formato de los ítems para la evaluación de capacidades fue definido por el equipo de trabajo, con base en estudios sobre el comportamiento de respuesta ante las tareas propuestas.

Las tareas en este caso, corresponden a los ítems o problemas cuya estructura ha sido pensada para favorecer cierta modalidad y magnitud de procesamiento cognitivo que permita obtener un resultado acorde con el objetivo de la medida. En este sentido, y dado el tenor del instrumento de evaluación aquí mencionado, se favoreció la construcción de ítems con los siguientes formatos:

- *Ítem de opción múltiple:* Consta de un enunciado y de tres o más opciones de respuesta, de la que sólo una es la correcta. Es el formato base, los restantes que aquí se presentan se pueden considerar variaciones de éste. Ejemplo:

La fórmula C_4H_{10} corresponde al siguiente hidrocarburo:

- a) Etano*
- b) Propano*
- c) Butano*

Este formato es de sobra conocido, sin embargo, no es tan familiar el hecho de que no es necesario contar con más de tres opciones de respuesta para contar con un buen ítem, siempre y cuando los distractores estén convenientemente contruidos (Abad, Olea y Ponsoda, 2001; Abad, Olea, Ponsoda y García, 2010).

A lo anterior, hay que sumar el que, contra la creencia habitual, que dicta que este tipo de ítems sólo sirve para evaluar tareas cognitivamente simples, lo cierto es que con ítems de elección múltiple bien contruidos es posible “evaluar tareas cognitivamente complejas, así como todos los pasos que deseen evaluarse de un proceso, como se pone de manifiesto ampliamente en la práctica de la evaluación” (Moreno, Martínez y Muñiz, 2006, p. 491).

- *Ítems de opción múltiple complejo:* Es necesario elegir una opción correcta de una serie referida a un conjunto previo de opciones. En general se presentan, como se ha sugerido, sólo tres opciones de respuesta, pero la elección de la alternativa correcta no se remite sólo a una de ellas, pudiendo el examinado seleccionar combinaciones posibles de dichas opciones. Ejemplo:

- I. Oviedo*
 - II. Pamplona*
 - III. Vigo*
-
- a) I y II*
 - b) II y III*
 - c) I, II y III*

Conviene preparar un ítem de esta clase, con tres opciones iniciales. Con más opciones, la dificultad del ítem no mejorará sustancialmente. Es mejor elaborar tres alternativas cuya combinación sea plausible y creíble, antes que aumentar el número de posibilidades.

- *Ítems con dependencia de contexto o multi-ítem:* Este tipo de ítems se construyen a partir de un problema o contexto inicial. Se propone una situación en la que se expone con algún detalle un problema frente al cual es necesario efectuar una lectura, análisis y reflexión detallada para responder posteriormente un conjunto de preguntas que están vinculadas a ese contexto. Esto genera cierta dependencia entre los ítems pues aluden al mismo contenido, razón por la cual y dado que se viola el supuesto de independencia de los ítems (que no haya contenido compartido entre ellos que ayude a dar pistas unos de otros), el contexto y sus ítems derivados son llamados testlets (Wainer y Lewis, 1990). Un ejemplo de este tipo de ítem son los clásicos encabezados que proponen un problema inicial y un conjunto de ítems que interrogan sobre dicho encabezado.

Para que la construcción del ítem sea aceptable, fue necesario respetar algunas directrices básicas, a saber: a) Mantener la homogeneidad de contenido, no variar el contenido ni el largo entre los distractores y respuesta correcta, pues ello tiene efectos en la discriminación, b) la inclusión de la opción “Ninguna de las anteriores”, aumenta innecesariamente la dificultad del ítem, c) por el contrario, la opción “Todas las anteriores” parece no tener mayor efecto en la dificultad del ítem, pero se sabe que afecta particularmente a los menos capaces y d) sumo cuidado en la redacción del encabezado del ítem pues debe comprenderse sin acceder a información adicional que no esté contenida en el contenido central del ítem (Martínez, Moreno, Martín y Trigo, 2009).

A las recomendaciones anteriores, se le sumaron otras sugeridas que, en muchos casos, cada disciplina respetó y puso en práctica en sus construcciones: a) Clasificación del proceso mental o cognitivo específico que se desea gatillar (interpretar, recordar, memorizar, etc.), b) contenido de cada ítem independiente del de otros ítems del test, salvo en los “testlets” donde ello no es posible, c) uso de un vocabulario sencillo para la redacción, d) contenido esencial y crucial en el encabezado, no en las opciones, e) evitar los dobles negativos como por ejemplo “no es posible no comunicar...” pues su lectura es confusa, f) todos los distractores son plausibles.

Una vez construidas las pruebas, se sometieron a doble análisis, tanto del equipo de trabajo como de ayudantes que revisaron aspectos formales de su desarrollo.

Construcción del instrumento de medición de capacidades pedagógicas.

Tal como se ha descrito, la prueba de capacidades pedagógicas es una evaluación transversal, aplicada a todos los estudiantes por igual, independiente de la disciplina a la que pertenecen, consta de: a) Un primer apartado de opción múltiple destinado a medir desempeños y b) un segundo apartado de escala tipo Likert destinado a medir actitudes hacia la docencia.

Tabla 2. Capacidades evaluadas por los responsables de cada unidad académica.

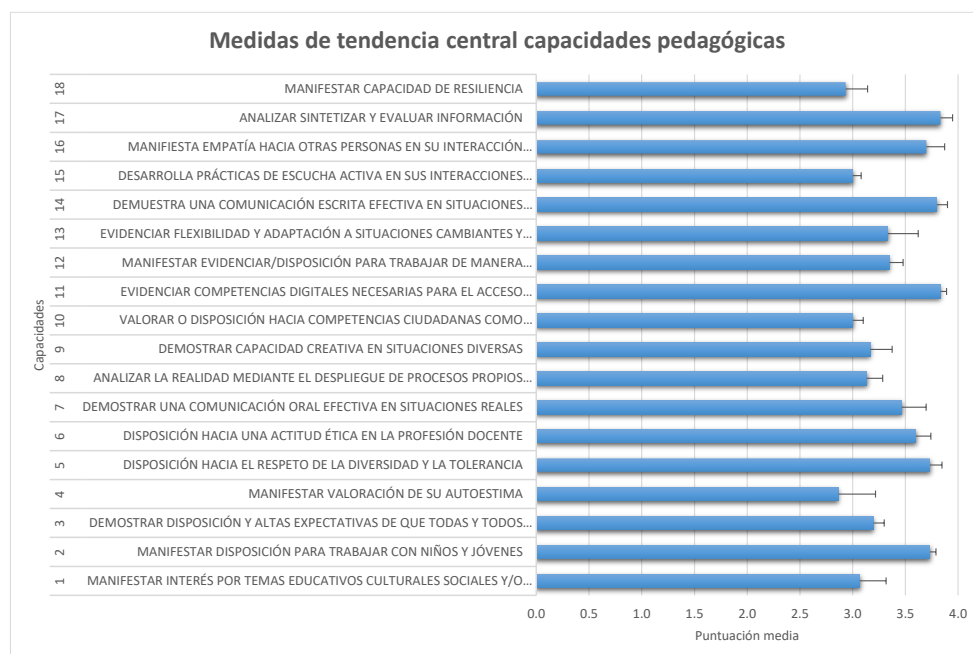
Id	Capacidad	Media	DE	Mínimo	Máximo
1	Manifestar interés por temas educativos culturales sociales y/o artísticos de nuestra realidad	3.1	0.25	2.8	3.3
2	Manifestar disposición para trabajar con niños y jóvenes	3.7	0.06	3.7	3.8
3	Demostrar disposición y altas expectativas de que todas y todos pueden aprender	3.2	0.10	3.1	3.3
4	Manifestar valoración de su autoestima	2.9	0.35	2.5	3.2
5	Disposición hacia el respeto de la diversidad y la tolerancia	3.7	0.12	3.6	3.8
6	Disposición hacia una actitud ética en la profesión docente	3.6	0.14	3.5	3.8
7	Demostrar una comunicación oral efectiva en situaciones reales	3.5	0.23	3.2	3.6
8	Analizar la realidad mediante el despliegue de procesos propios de un pensamiento crítico	3.1	0.15	3	3.3
9	Demostrar capacidad creativa en situaciones diversas	3.2	0.21	3	3.4
10	Valorar o disposición hacia competencias ciudadanas como necesarias para actuar de manera constructiva en el entorno socioeducativo	3.0	0.10	2.9	3.1
11	Evidenciar competencias digitales necesarias para el acceso gestión y producción de información	3.8	0.06	3.8	3.9
12	Manifestar evidenciar/disposición para trabajar de manera autónoma cooperativa y colaborativa	3.4	0.13	3.2	3.5
13	Evidenciar flexibilidad y adaptación a situaciones cambiantes y difíciles	3.3	0.29	3	3.5
14	Demuestra una comunicación escrita efectiva en situaciones reales	3.8	0.10	3.7	3.9
15	Desarrolla prácticas de escucha activa en sus interacciones habituales	3.0	0.08	2.9	3.1
16	Manifiesta empatía hacia otras personas en su interacción habitual	3.7	0.17	3.5	3.8
17	Analizar sintetizar y evaluar información	3.8	0.12	3.7	3.9
18	Manifestar capacidad de resiliencia	2.9	0.21	2.7	3.1
Total		3.4	0.36	2.5	3.9

El primer apartado de la prueba, aquel de capacidades transversales, fue elaborado a través de consulta a los responsables de todas las unidades académicas implicadas. La intención era que validaran un conjunto de atributos (ver tabla 2), conceptualizados como capacidades, que mediante una extensiva revisión de literatura, configuraron una lista de 18 opciones. A los responsables se les pidió que se pronunciaron por la pertinencia de la capacidad, la prioridad dada a cada una de ellas y su factibilidad.

Cada capacidad entonces fue evaluada por los participantes en estos tres aspectos, seleccionado una opción de respuesta de una escala de Acuerdo-Desacuerdo desde 1 a 4 puntos. Mayor valor implicó mayor relevancia dada a lo que se desea evaluar. En la tabla 2 se disponen los estadísticos obtenidos a partir de esta escala de Acuerdo-Desacuerdo. En primer lugar se calculó la media y desviación típica de cada una de las capacidades pero en términos generales, sin separar por valoración de enunciado, prioridad o factibilidad.

Como se aprecia en la tabla 2, las capacidades con mayor grado de acuerdo para conformar un conjunto de aptitudes a evaluar en el perfil de ingreso, corresponden a: a) Manifestar disposición para trabajar con niños y jóvenes (2), b) disposición hacia el respeto de la diversidad y la tolerancia (5), c) evidenciar competencias digitales necesarias para el acceso gestión y producción de información (11), d) demuestra una comunicación escrita efectiva en situaciones reales (14), e) manifiesta empatía hacia otras personas en su interacción habitual (16) y f) analizar sintetizar y evaluar información (17). Las medias de estas capacidades están tres o cuatro décimas por sobre la media general de capacidades.

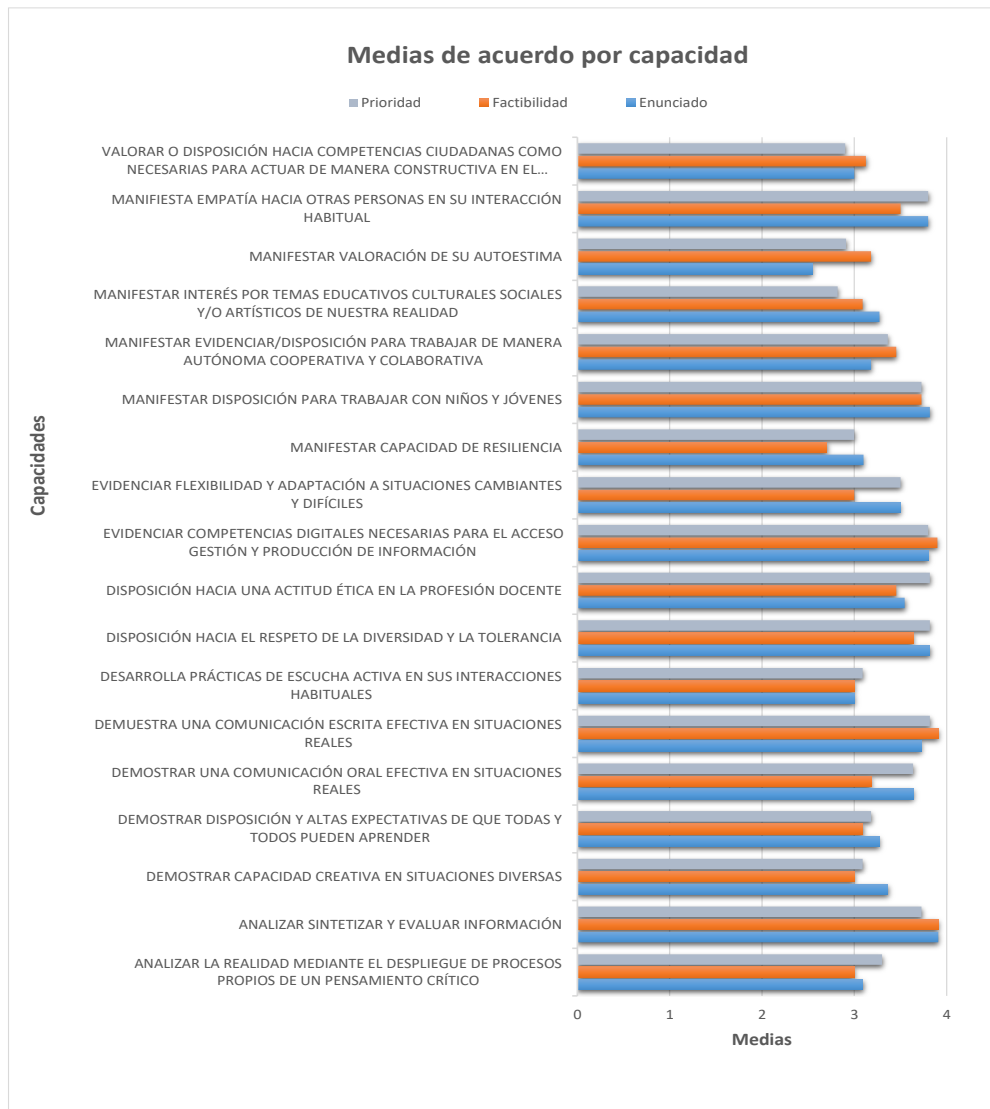
Figura 1. Medias y desviaciones estándar por capacidad.



Se puede observar el mismo resultado mediante un gráfico general de barras (Figura 1). Las barras corresponden a las medias observadas anteriormente, se incluyen las desviaciones estándar mediante las líneas de error en la parte superior.

En la figura 1 se observa claramente cuáles son las capacidades que concitan mayor acuerdo y coinciden con las que ya se han mencionado. En consecuencia, de las 18 capacidades originales, es posible privilegiar las seis ya mencionadas, que podrán eventualmente componer la estructura de evaluación en el ámbito pedagógico.

Figura 2. Medias por dimensión de cada capacidad.



Un análisis más detallado de esta información considera además el acuerdo con la estructura que enuncia la capacidad (Enunciado), la prioridad dada a dicha capacidad (Prioridad) y la factibilidad de implementarlo (Factibilidad). El resultado se dispone en la precedente figura 2.

Los resultados obtenidos segregando por los atributos evaluados siguen favoreciendo a las capacidades que han sido resaltadas anteriormente, pero con matices:

a) En disposición para trabajar con niños y jóvenes hay consenso en la relevancia de los tres aspectos considerados, estructura del enunciado, de su prioridad y factibilidad de evaluación; b) En disposición hacia el respeto de la diversidad y la tolerancia también hay un acuerdo amplio hacia su inclusión en la evaluación, pero se le ve menos factible de evaluar, lo que se explica porque se trata de actitudes; c) En evidenciar competencias digitales necesarias para la gestión y producción de información, su valoración es alta y se le considera muy factible de evaluar; d) en la demostración de una comunicación escrita efectiva en situaciones reales, se le considera factible de evaluar (hay varias pruebas de comprensión y producción de textos), pero el conceptualización de la capacidad concita menos acuerdo; e) en la manifestación de empatía hacia las personas, se le considera menos factible de evaluar en relación a sus otros atributos, lo que también remite a la cualidad del atributo medido; f) Analizar, sintetizar y evaluar información representa también una capacidad de relevancia, pero su prioridad denota menor acuerdo.

Tabla 3. Porcentaje de acuerdo según opciones de la escala de evaluación.

Jueces	Bajo acuerdo (1)		Acuerdo Regular (2)		Acuerdo aceptable (3)		Alto acuerdo (4)	
	f	%	f	%	f	%	f	%
Juez 1	0	0.0	3	8.1	10	27.0	24	64.9
Juez 2	10	27.0	4	10.8	14	37.8	9	24.3
Juez 3	0	0.0	1	2.7	10	27.0	26	70.3
Juez 4	5	13.5	3	8.1	6	16.2	23	62.2
Juez 5	2	5.4	4	10.8	6	16.2	25	67.6
Juez 6	4	10.8	2	5.4	14	37.8	17	46.0
Juez 7	0	0.0	2	5.4	7	18.9	28	75.7
Juez 8	0	0.0	6	16.2	13	35.1	18	48.7
Juez 9	0	0.0	0	0.0	16	43.2	21	56.8
Juez 10	0	0.0	6	16.2	15	40.5	16	43.2
Juez 11	0	0.0	1	2.7	6	16.2	30	81.1
Total	21	5.2	32	7.9	117	28.8	237	58.2

Considerando cada uno de los responsables de las unidades académicas, esto es, jefes de carrera y didactas, como jueces en el proceso de selección de las capacidades, es posible obtener alguna medida del acuerdo de sus juicios para detectar si las evaluaciones por ellos establecidas son fiables. Los resultados se disponen en la tabla 3.

Se observa que la mayoría de los jueces emiten juicios hacia el alto acuerdo, pues los mayores porcentajes de elección se concentran en las dos opciones de respuesta más altos. La opción 4, salvo el juez 2 y 10, siempre acapara al menos el 50% de las elecciones realizadas por los jueces.

Es posible aplicar alguna medida de consistencia para determinar la fiabilidad del juicio de los académicos que actuaron como jueces. Aunque hay varias medidas, una posibilidad es aplicar el coeficiente alfa de Cronbach, usualmente utilizado para determinar consistencia entre ítems. No evalúa concretamente concordancia, pero brinda una medida de consistencia interna, que sugiere si los juicios de los académicos fueron coherentes en mayor o menor medida.

Para aplicar alfa, en este caso fue necesario calcular las correlaciones policóricas entre juicios, pues sus distribuciones, como se aprecia en la tabla 3, están claramente sesgadas hacia los valores altos de la escala. Con dichas correlaciones, se aplicó alfa de Cronbach. El resultado obtenido fue igual a 0.74 (la escala va de 0 a 1), lo que indica que el grado de fiabilidad de los juicios en conjunto es aceptable.

La sección destinada a escala de actitudes, con formato tipo Likert, fue construida por el equipo asesor. Se determinó que algunas de las dimensiones anteriores no podrían ser evaluadas desde una perspectiva de logro pero sí apuntaban a aspectos actitudinales y por tanto lo que correspondía era medirlas mediante las tareas propuestas para evaluación de actitudes. La escala asociada fue de 5 puntos, desde Muy en Desacuerdo (1) a Muy de Acuerdo (5).

La tabla de especificaciones original dio origen a una tabla reducida. Las dimensiones establecidas en dicha tabla, y que permitió construir la prueba de capacidades pedagógica, se disponen en la tabla siguiente:

Tabla 4. Dimensiones test de capacidades pedagógicas.

Código	Dimensión	Ítems	Rasgo medido
CP1	Manifestar disposición para trabajar con niños y jóvenes.	1 al 5	Actitud
CP2	Disposición hacia el respeto de la Diversidad y la Tolerancia.	6 al 10	Actitud
CP3	Manifiesta Empatía hacia otras personas en su interacción habitual.	11 al 15	Actitud
CP4	Evidenciar reflexión crítica y propositiva.	16 al 20	Desempeño
CP5	Evidenciar Competencias Digitales necesarias para el acceso, gestión y producción de información.	21 al 26	Desempeño
CP6	Conocimiento Metacognitivo acerca de la Producción de Textos Escritos	27 al 29	Desempeño
CP7	Comprensión de Textos Escritos	30 al 34	Desempeño

En definitiva, el instrumento completo contó con 15 afirmaciones en formato escala, y 20 preguntas en diferentes formatos de opción múltiple, para un total de 34 ítems.

Construcción del instrumento de medición de capacidades disciplinares.

La construcción de las distintas versiones de la prueba de capacidades disciplinares estuvo a cargo de equipos conformados por las unidades académicas correspondientes. En consecuencia, fueron ellos los que definieron las dimensiones a la base de los instrumentos. No obstante tenían la guía del tipo de ítems que debían construir y por tanto, los formatos aceptados para aunar el tipo de instrumento a aplicar.

Se pidió a las unidades académicas que crearan sus propias matrices de especificaciones y definieran las capacidades propias de la carrera. Sólo se solicitó que intentaran construir dicha matriz con la siguiente estructura: a) Definición de capacidad, b) Procesos de pensamiento, c) Habilidad, d) Indicadores, e) Ítems y f) Ponderación en la prueba.

El número de ítems de cada prueba fue variable. Aunque se pidió un total de 60 ítems, para contar con un mayor número de ítems para elegir. Posteriormente en las pruebas finales se incluyeron poco menos de la mitad de éstas, pues de lo contrario, el instrumento hubiese resultado demasiado largo. La tabla 5 dispone el número de ítems disciplinares por carrera.

Tabla 5. Número de ítems por test en cada unidad académica.

Tests capacidades disciplinares	Número de preguntas
Biología	26
Educación Básica	29
Educación Diferencial	29
Educación Física	29
Educación Parvularia	29
Filosofía	29
Física	27
Historia	29
Inglés	29
Literatura y Ciencias Lenguaje	29
Matemáticas	29
Música	28
Química	28

En las carreras se ha respetado el número de ítems sugerido. En definitiva, los estudiantes debieron responder un instrumento compuesto por 15 ítems tipo Likert, 20 ítems de capacidades transversales y entre 26 y 29 ítems de capacidades disciplinares.

Procedimiento

El procedimiento de elaboración comenzó durante abril, con la puesta en ejercicio del equipo asesor que planteó una modalidad de evaluación viable para las 13 carreras. Después de varias reuniones en las que participaron profesionales de la Escuela de Pedagogía, se separó la evaluación en aquellas capacidades transversales de índole pedagógica, de aquellas en las que era necesario el concurso de los profesionales de cada carrera y que en definitiva determinaron las capacidades propias de cada disciplina que debían ser evaluadas. A partir de este punto, el proceso de trabajo implicó:

1. Propuesta de un listado de capacidades transversales que fueron sometidas a consulta entre los profesionales de las unidades académicas comprometidas. El resultado de la elección realizada por éstos, se discutió en el apartado instrumentos.
2. Indicación del formato y estructura de los instrumentos que debían ser desarrollados por las diferentes carreras.
3. Solicitud a las unidades académicas, para que elaboraran sus propias capacidades, y construyeran ítems en función de dichas aptitudes. Se sugirió un formato de presentación para la construcción de las tablas de especificaciones.
4. Devolución de las tablas de especificaciones de las carreras con el número solicitado de ítems al equipo asesor.
5. Revisión formal por parte de ayudantes para determinar si los ítems cumplen con el formato solicitado.
6. Segunda revisión, esta vez sobre el contenido, realizado por los miembros del equipo asesor, para decidir sobre ítems adecuados o no.
7. Recopilación de todo el material disponible y aprobado y construcción de la prueba de capacidades transversales y específicas.
8. Implementación de las pruebas en la plataforma de encuestas online LimeSurvey (<https://www.limesurvey.org/>). Revisión y corrección de la versión en línea.

9. Recolección presencial de datos mediante tabletas electrónicas facilitadas a los estudiantes en sus respectivas aulas, asegurando una tasa mínima de respuesta, que fue cercana al 90%.
10. Recopilación de la información y envío digital para su análisis.
11. Análisis de datos y emisión de informe psicométrico.
12. Análisis de datos e informe de resultados por carrera.

En este informe, se da cuenta del último de los procesos mencionados, abordando los resultados por carrera.

Análisis

El análisis implicó en primer lugar la descripción de los datos de medidas de tendencia central y variabilidad, mínimo y máximo para cada ítem y dimensión de las pruebas y segregando por cada carrera. Se adjuntó el cálculo de las correlaciones biserial-puntuales para dar cuenta del nivel de discriminación (Martínez, Hernández y Hernández, 2006).

Una vez estructuradas las dimensiones de modo empírico y reducido el número de ítems, conforme su dificultad y discriminación, se obtiene la fiabilidad mediante coeficiente alfa de Cronbach para cada una de dimensiones (Santisteban, 2009).

En el apartado de resultados, en la sección actitudinal, sólo se informan descriptivos para las puntuaciones en la escala de 1 a 5 (Muy en desacuerdo a Muy de acuerdo), en la sección capacidades pedagógicas y capacidades disciplinares, se disponen los mismos análisis pero en función de porcentajes de logro, para una lectura más sencilla.

Para llevar a cabo los análisis se utilizó el entorno de programación R (Ihaka y Gentleman, 1996; R Development Core Team, 2017) y la hoja de cálculo Calc de OpenOffice versión 3.5. El producto de este proceder se presenta con mayor detalle en la discusión de este estudio, en que se afianzan las características y dificultades mostradas por los instrumentos.

Resultados

Composición de los instrumentos.

De la prueba pedagógica.

Una vez realizados los análisis de cada uno de los ítems de la propuesta de prueba para capacidades y actitudes docentes transversales, los ítems que se mantuvieron en cada apartado se disponen a continuación.

En el caso de la escala de actitudes asociada con la evaluación, se muestra en esta tabla sólo los ítems que deben ser conservados en función de sus propiedades formales, con objeto de mejorar el desempeño general del instrumento.

Tabla 6. Descriptivos para ítems de la escala de actitudes de la prueba pedagógica.

Ítems	Encabezados	Todas las carreras (n=623)					
		Media	DE	Mediana	Asimetría	Curtosis	Discriminación
SQ01	En general, pienso que el trato	3.96	0.91	4	-0.71	0.29	0.26
SQ02	Cada niño es un “mundo”	4.52	0.85	5	-1.9	3.21	0.28
SQ03	Con los niños es importante servir	4.67	0.66	5	-2.43	7.2	0.22
SQ04	La cualidad por excelencia de	4.53	0.66	5	-1.26	1.26	0.22
SQ06	Por principio, todos los seres humanos	4.76	0.6	5	-2.78	8.14	0.31
SQ10	La sociedad debiera establecer	4.79	0.55	5	-3.17	11.97	0.28
SQ11	Cuando alguien está sufriendo	3.52	1	4	-0.4	-0.14	0.28
SQ12	Cuando hablo con alguien que está	4.18	0.95	4	-1.2	1.26	0.30
SQ14	Cuando veo que se burlan de una persona	3.63	1.19	4	-0.63	-0.46	0.35
SQ15	Cuando tengo una dificultad con alguien	4.26	0.81	4	-0.9	0.38	0.21

Se han eliminado de la tabla los ítems 5, 7, 8, 9 y 13. Esto es porque algunos de los valores de discriminación estaban fuera de lo esperado.

En cuanto a los ítems del apartado de capacidades pedagógicas, correspondiente a la evaluación del desempeño de los estudiantes, la tabla siguiente dispone los resultados definitivos, habiendo eliminado los elementos que no tuvieron un buen comportamiento. Se dispone en la tabla 7.

Tabla 7. Descriptivos para ítems de la escala de desempeños de la prueba pedagógica.

Ítems	Encabezado	Todas las carreras (n=623)					
		Media	DE	Mediana	Asimetría	Curtosis	Discriminación
P16	De acuerdo al Documento 1, el país	0.96	0.21	1	-4.38	17.23	0.16
P18	De acuerdo a los porcentajes (%) de Basándose en el Documento 2, en el caso de los países	0.97	0.16	1	-5.98	33.85	0.18
P20	De acuerdo al Documento 4: Juan Pablo debe buscar información sobre	0.78	0.42	1	-1.33	-0.24	0.16
P22	De acuerdo al Documento 5: Google Académico	0.97	0.17	1	-5.61	29.54	0.28
P23	Basándose en el Documento 6: Natalia decidió que el cálculo de la Nota Final	0.65	0.48	1	-0.62	-1.62	0.14
P25	El Plagio de Información puede ser definido como el uso de ideas y/o producciones	0.8	0.4	1	-1.47	0.15	0.14
P26	Cuando termino de escribir un trabajo que me dan en la Universidad, lo primero que hago es:	0.62	0.49	1	-0.51	-1.75	0.13
P27	De acuerdo a lo leído en el Documento 7, en una escuela rural multigrado unitaria, la educadora de párvulos	0.82	0.39	1	-1.64	0.68	0.07
P31	La expresión “Esta es la realidad” marcada en el párrafo 3, del Documento 8, se refiere a:	0.62	0.49	1	-0.5	-1.75	0.11
P32		0.84	0.37	1	-1.85	1.41	0.12

Los ítems de la parte referida al desempeño de los estudiantes se reducen en función de los resultados que se observan anteriormente en la tabla 7. Se removieron los ítems 17, 19, 21, 24, 28, 29, 30, 33 y 34, para mostrar la composición ideal del instrumento como debiera ser.

De las pruebas disciplinares.

Del mismo modo que en el caso del instrumento pedagógico transversal, las pruebas desarrolladas por cada una de las unidades académicas fueron reducidas en función del comportamiento de los ítems en dificultad y discriminación. Las tablas que se muestran a continuación, sólo disponen los ítems que fueron utilizados para calcular los resultados finales que se disponen en la segunda parte de este informe.

Su presentación, que no es otra que la dispuesta en el informe uno, pero removiendo todos los ítems que no fueron utilizados para calcular los resultados finales de cada estudiante. En consecuencia, las tablas disponen los mejores ítems que son justamente aquellos que han servido para obtener resultados definitivos.

El análisis formal que acompaña cada ítem, implica la dificultad (media para variables dicotómicas), la desviación estándar o típica y el índice de discriminación calculado mediante la aproximación de correlación ítem-total.

a) *Carrera de Biología.*

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 8.

Tabla 8. Dificultad y discriminación para la carrera de Biología.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P35	Las bacterias que viven en nuestra boca	0.74	0.44	0.51
P36	Se hizo un experimento con “dientes de leche”	0.71	0.46	0.35
P37	El espinoso es un pez	0.68	0.48	0.07
P38	Un cocinero hace el pan	0.29	0.46	0.45
P40	Es fácil encontrar en la publicidad	0.74	0.44	0.28
P41	Los progresos en el diagnóstico	0.61	0.50	0.36
P42	El siguiente gráfico muestra	0.03	0.18	0.26
P44	El yogurt tuvo sus orígenes	0.68	0.48	0.17
P47	El nivel de dióxido de carbono	0.45	0.51	0.27
P50	Estudios sobre la densidad promedio	0.13	0.34	0.34
P51	De acuerdo a recientes estadísticas	0.84	0.37	0.34
P52	Basándose en la información anterior,	0.16	0.37	0.34
P54	Los grupos ecologistas exigen	0.58	0.50	0.26
P56	El almidón es un polisacárido	0.13	0.34	0.09
P57	A partir del análisis de la siguiente	0.74	0.44	0.28
P58	La densidad del agua es	0.32	0.48	0.32
P59	La figura muestra la prevalencia	0.61	0.50	0.02
P60	A partir de la imagen anterior	0.65	0.49	0.39
P61	Francia es un país que históricamente	0.58	0.50	0.23

Con base en la remoción, eliminados los ítems 39, 43, 45, 46, 48, 49, 53 y 55, se obtuvo esta prueba resultante.

b) Carrera de Educación Básica

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 9:

Tabla 9. Dificultad y discriminación para la carrera de Educación Básica.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P36	A partir de la lectura del Documento 1, ¿cuál de las siguientes	0.51	0.51	0.19
P39	La finalidad del afiche publicitario	0.79	0.41	0.42
P40	Como efectos de la mala nutrición	0.86	0.35	0.11
P45	El problema que plantea la lectura en	0.49	0.51	0.21
P46	En una Escuela hay 600 alumnos	0.19	0.39	0.02
P49	Lee atentamente la siguiente conversación	0.37	0.49	0.26
P55	Del Documento 4 se puede desprender	0.67	0.47	0.16
P56	A partir de la información de la infografía	0.79	0.41	0.44
P57	¿Cuál de los siguientes enunciados	0.72	0.45	0.10
P59	De acuerdo a la información proporcionada	0.77	0.43	0.13
P60	Una educadora de párvulos de un jardín	0.60	0.49	0.14
P61	La directiva de apoderados del	0.91	0.29	0.11
P62	La profesora jefe de Quinto Básico	0.51	0.51	0.39
P63	La educadora de párvulos del	0.23	0.43	0.14
P64	La profesora jefe del Sexto Básico de	0.47	0.50	0.56

Con base en la remoción, eliminados los ítems 35, 37, 38, 41, 42, 43, 44, 47, 48, 50, 51, 52, 53, 54 y 58, se obtuvo esta prueba resultante.

c) Carrera de Educación Diferencial:

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 10:

Tabla 10. Dificultad y discriminación para la carrera de Educación Diferencial.

Ítems	Encabezado	Dificultad	Desv Estándar.	Discriminación
P35	A partir de la lectura del Documento 1, ¿cuál de	0.83	0.38	0.41
P36	A partir de la lectura del Documento 1, ¿cuál de las siguientes	0.43	0.50	0.29
P38	A partir del cuadro anterior, clasifica	0.97	0.18	0.30
P39	La finalidad del afiche publicitario	0.71	0.46	0.16
P41	Respecto del de logro de los estudiantes	0.16	0.37	0.18
P42	Con base en la información del Documento	0.16	0.37	0.24
P43	A partir del Documento 2, para poder	0.03	0.18	0.21

P45	El problema que plantea la lectura en	0.48	0.50	0.00
P47	Rocío es una Educadora Diferencial recién egresada	0.45	0.50	0.03
P49	Lee atentamente la siguiente conversación	0.53	0.50	0.23
P54	El titular más apropiado para la	0.53	0.50	0.19
P55	Del Documento 4 se puede desprender	0.67	0.47	0.09
P57	¿Cuál de los siguientes enunciados	0.66	0.48	0.13
P58	¿Qué alternativa es correcta de acuerdo	0.05	0.22	0.12
P60	Una educadora de párvulos de un jardín	0.67	0.47	0.51
P61	La directiva de apoderados del	0.81	0.40	0.03
P62	La profesora jefe de Quinto Básico	0.52	0.50	0.20
P64	La profesora jefe del Sexto Básico de	0.28	0.45	0.10

Se eliminaron los ítems: 37, 40, 44, 46, 48, 50, 51, 52, 53, 56, 59 y 63. Con base en esta información se obtuvo la prueba definitiva.

d) Carrera de Educación Física:

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 11:

Tabla 11. Dificultad y discriminación para la carrera de Educación Física.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P35	A partir de la Figura 1, identifica	0.95	0.21	0.22
P36	Los deportes individuales o de conjunto	0.92	0.28	0.05
P37	¿Cuál de las imágenes de la Figura 2	0.34	0.48	0.06
P39	En la ejecución de todos los deportes	0.95	0.21	0.07
P42	Realizar actividad física moderada	0.26	0.44	0.18
P45	A partir de la información del siguiente gráfico,	0.91	0.29	0.09
P49	Considerando el Documento 2	0.27	0.45	0.16
P50	De las palabras del barón de Pierre	0.48	0.50	0.24
P53	De acuerdo al Documento 4, objetivo estratégico	0.58	0.50	0.08
P54	La política nacional del deporte	0.41	0.50	0.03
P56	Señala la alternativa que, en su conjunto	0.13	0.34	0.05
P58	Según el siguiente gráfico, selecciona	0.95	0.21	0.09
P61	Considerando el Documento 6, es correcto	0.68	0.47	0.13
P63	Considerando el gráfico anterior, ¿cuáles	0.39	0.49	0.06

Se eliminaron los ítems: 38, 40, 41, 43, 44, 46, 47, 48, 51, 52, 55, 57, 59, 60, 62 y 64.

e) *Carrera de Educación Parvularia:*

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 12:

Tabla 12. Dificultad y discriminación para la carrera de Educación Parvularia.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P35	A partir de la lectura del Documento 1, ¿cuál de	0.79	0.41	0.04
P36	A partir de la lectura del Documento 1, ¿cuál de las siguientes	0.38	0.49	0.06
P39	La finalidad del afiche publicitario	0.77	0.43	0.29
P44	El título del Documento 3 podría	0.95	0.22	0.01
P45	El problema que plantea la lectura en	0.44	0.50	0.15
P47	Rocío es una Educadora Diferencial recién egresada	0.33	0.48	0.03
P48	Observa las siguientes pirámides:	0.08	0.27	0.03
P49	Lee atentamente la siguiente conversación	0.64	0.49	0.16
P55	Del Documento 4 se puede desprender	0.67	0.48	0.13
P56	A partir de la información de la infografía	0.74	0.44	0.31
P57	¿Cuál de los siguientes enunciados	0.67	0.48	0.13
P62	La profesora jefe de Quinto Básico	0.46	0.51	-0.02
P64	La profesora jefe del Sexto Básico de	0.56	0.50	0.19

Se eliminaron los ítems: 37, 38, 40, 41, 42, 43, 46, 50, 51, 53, 54, 58, 59, 60, 61 y 63. No se eliminó el ítem 62 a pesar de su baja discriminación, porque su dificultad era aceptable y porque el instrumento hubiese quedado muy reducido.

f) *Carrera de Filosofía:*

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 13:

Tabla 13. Dificultad y discriminación para la carrera de Filosofía.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P35	“Nadie se baña en el río dos veces	0.55	0.50	0.24
P39	Algunos de los aportes de René Descartes	0.32	0.47	0.27
P41	La ética tiene tres tareas fundamentales	0.27	0.45	0.24
P42	Según el Texto 1, Antístenes sostenía	0.63	0.49	0.23
P44	En el Texto 1 se recomienda	0.97	0.18	0.26
P45	En el Texto 1, la ironía de Antístenes	0.82	0.39	0.35
P46	¿Qué disciplinas se relacionan	0.88	0.32	0.26
P47	En todo movimiento o cambio, Aristóteles	0.27	0.45	0.13
P48	La fuerza de gravedad atrae	0.38	0.49	0.19

P51	Los anteriores planteamientos del Texto	0.40	0.49	0.23
P53	Los textos nos permiten:	0.57	0.50	0.27
P54	Los elementos simples de la materia	0.58	0.50	0.23
P55	Según la teoría tripartita de Platón	0.33	0.48	0.31
P56	Indique en el silogismo descrito	0.27	0.45	0.29
P57	Indique cuál es el término	0.18	0.39	0.39
P58	Indique cuál es el término medio	0.87	0.34	0.33
P59	Afirmar que “el Rey de España es	0.95	0.22	0.32
P61	Según el Texto 4, para Tomás	0.50	0.50	0.29
P63	Según el Texto 5, es posible	0.45	0.50	0.17
P64	De acuerdo al Texto 5, los seres son:	0.60	0.49	0.32

Se eliminaron los ítems: 36, 37, 38, 40, 43, 49, 50, 52, 60 y 62.

g) Carrera de Física:

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 14:

Tabla 14. Dificultad y discriminación para la carrera de Física.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P35	El gráfico presentado a continuación	0.83	0.38	0.49
P37	Un autobús circula por un tramo recto	0.80	0.40	0.18
P38	“Las variaciones en la distancia de la Tierra	0.80	0.40	0.34
P39	La figura muestra dos esferas con carga	0.29	0.46	0.36
P42	Cuando un paracaidista salta desde	0.85	0.36	0.16
P43	De acuerdo con la lectura del Documento 1	0.85	0.36	0.28
P45	Considerando el movimiento del	0.85	0.36	0.30
P46	El efecto de la lluvia ácida en el mármol	0.68	0.47	0.37
P47	La luz verde de un semáforo se	0.54	0.50	0.18
P48	El péndulo de la figura oscila, tardando	0.49	0.51	0.18
P51	En un puesto fronterizo, el control	0.76	0.43	0.29
P52	En una carretera, cada 4,0 [km]	0.61	0.49	0.35
P55	Para observar estrellas de escaso brillo,	0.78	0.42	0.44
P57	A mayor fuerza del viento, las palas	0.71	0.46	0.29
P58	A igual velocidad del viento, si los	0.78	0.42	0.18
P62	Roberto hizo un experimento para investigar	0.80	0.40	0.38

Se eliminaron los ítems: 36, 40, 41, 44, 49, 50, 53, 54, 56, 59, 60 y 61.

h) Carrera de Historia:

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 15:

Tabla 15. Dificultad y discriminación para la carrera de Historia.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P37	¿Con qué imperio es posible	0.68	0.47	0.14
P38	Si quisiéramos determinar la suficiencia	0.61	0.49	0.18
P39	Al seguir una secuencia para establecer	0.77	0.42	0.15
P41	A partir del Documento 2 es posible	0.44	0.50	0.22
P42	Al seguir a una secuencia para comparar	0.72	0.45	0.19
P43	El Documento 3 es una fuente de tipo	0.62	0.49	0.10
P45	El Documento 4 es un extracto de un	0.56	0.50	0.29
P48	De acuerdo al Documento 5, sería correcto	0.42	0.50	0.12
P49	A partir del Documento 5 es posible inferir	0.54	0.50	0.12
P50	Considerando el Documento 6, sería correcto	0.37	0.49	0.26
P51	De acuerdo al documento 6, Giovanni Bocaccio	0.69	0.47	0.12
P52	Los siguientes pasos corresponden a una	0.63	0.49	0.14
P53	Al comparar las pirámides de población	0.79	0.41	0.23
P54	¿Cuál de las siguientes políticas públicas	0.69	0.47	0.12
P55	Desde la perspectiva de la institucionalidad	0.93	0.26	0.15
P57	En la Constitución Política de Chile	0.87	0.34	0.32
P59	Considerando el Documento 10, es posible	0.80	0.40	0.12
P61	Considerando el gráfico del Documento 11	0.20	0.40	0.18
P62	El Documento 11 presenta la situación de Fecundidad	0.83	0.38	0.19
P63	Si tuvieras que hacer una investigación	0.69	0.47	0.14
P64	En la ciudad de Viña del Mar	0.63	0.49	0.30

Ítems que fueron removidos: 35, 36, 40, 44, 46, 47, 56, 58 y 60. No obstante, toda la prueba refleja unas propiedades psicométricas muy pobres en cuanto el grado de discriminación. Se han dejado varios ítems sin embargo, pues la extensión del instrumento se vería severamente afectada.

i) Carrera de Inglés:

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 16:

Tabla 16. Dificultad y discriminación para la carrera de Inglés.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P35	En el párrafo 1 del Texto 1, el autor	0.86	0.35	0.30
P36	En el párrafo 3 del Texto 1, la mención a	0.70	0.46	0.22
P37	De acuerdo al contenido del Texto 1, un	0.68	0.47	0.26
P38	El propósito del Texto 1 es:	0.50	0.50	0.12
P39	De acuerdo al Texto 1, lingüística se define	0.64	0.48	0.17
P40	De acuerdo al Texto 1, la palabra prominentes	0.56	0.50	0.18
P41	¿A qué tipo de texto corresponde el	0.63	0.49	0.28
P42	De acuerdo al Texto 1, la palabra rótulo	0.58	0.50	0.43
P44	La función del gráfico del Texto 2 es:	0.76	0.43	0.27
P46	Los distintos colores que marcan zonas en el	0.71	0.46	0.37
P47	En la infografía del Texto 3, el término	0.41	0.50	0.15
P48	El propósito comunicativo de la infografía	0.53	0.50	0.13
P49	El título que mejor resume el contenido	0.85	0.36	0.18
P50	Según el Texto 4, las alergias son causadas	0.49	0.50	0.22
P51	De acuerdo al Texto 4, Alérgeno se	0.69	0.47	0.12
P55	De acuerdo al Texto 5, se puede caracterizar	0.84	0.37	0.15
P56	El contenido del Texto 5 se organiza mediante	0.70	0.46	0.35
P57	De acuerdo al Texto 6, ¿por qué los habitantes	0.96	0.19	0.39
P58	En el Texto 6, ¿por qué las ciudades de	0.54	0.50	0.32
P59	Según el Texto 7, ¿por qué el trabajo de	0.85	0.36	0.27
P62	Lea el siguiente texto y luego responda la pregunta	0.64	0.48	0.21
P63	Lea el siguiente enunciado y responda la pregunta	0.70	0.46	0.22

Ítems que fueron eliminados: 43, 45, 52, 53, 54, 60, 61 y 64.

j) *Carrera de Literatura y Ciencias del Lenguaje:*

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 17:

Tabla 17. Dificultad y discriminación para la carrera de Literatura y Ciencias del Lenguaje.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P38	El propósito del Texto 1 es:	0.76	0.43	0.10
P39	De acuerdo al Texto 1, lingüística se define	0.76	0.43	0.13
P42	De acuerdo al Texto 1, la palabra rótulo	0.84	0.37	0.15
P46	Los distintos colores que marcan zonas en el	0.84	0.37	0.22
P48	El propósito comunicativo de la infografía	0.43	0.50	0.25
P51	De acuerdo al Texto 4, Alérgeno se	0.78	0.42	0.16
P54	Las categorías que sirven de base para	0.41	0.50	0.12

P56	El contenido del Texto 5 se organiza mediante	0.89	0.32	0.10
P59	Según el Texto 7, ¿por qué el trabajo de	0.84	0.37	0.26
P62	Lea el siguiente texto y luego responda la pregunta	0.43	0.50	0.16

Ítems que fueron eliminados producto de su dificultad y discriminación: 35, 36, 37, 40, 41, 43, 44, 45, 47, 49, 50, 52, 53, 55, 57, 58, 60, 61, 63 y 64.

k) Carrera de Matemáticas:

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 18:

Tabla 18. Dificultad y discriminación para la carrera de Matemáticas.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P35	Dos números enteros se dicen	0.55	0.50	0.24
P39	¿Cuál de las siguientes afirmaciones	0.32	0.47	0.27
P41	Se dice que un número natural	0.27	0.45	0.24
P42	La siguiente tabla corresponde	0.63	0.49	0.23
P44	De acuerdo al Documento 1, después	0.97	0.18	0.26
P45	De acuerdo al Documento 1, ¿en qué	0.82	0.39	0.35
P46	De acuerdo al Documento 1, ¿cuántas	0.88	0.32	0.26
P47	¿Cuál de las siguientes expresiones	0.27	0.45	0.13
P48	¿Cuál es la dosis para una persona	0.38	0.49	0.19
P51	De acuerdo al Documento 2, ¿para cuántas	0.40	0.49	0.23
P53	Con base en el Documento 3, si el promedio	0.57	0.50	0.27
P54	De acuerdo a las opciones de sueldo	0.58	0.50	0.23
P55	Una nave espacial viaja de la	0.33	0.48	0.31
P56	Una pizzería ofrece pizzas	0.27	0.45	0.29
P57	¿En qué periodo se presenta el	0.18	0.39	0.39
P58	De acuerdo al Documento 4, ¿qué	0.87	0.34	0.33
P59	De acuerdo al Documento 4, ¿qué región	0.95	0.22	0.32
P61	Si $x \in Z$, se puede	0.50	0.50	0.29
P63	De acuerdo al Documento 5, si $T k$	0.45	0.50	0.17
P64	De acuerdo al Documento 5	0.60	0.49	0.32

Ítems que fueron eliminados producto de su dificultad y discriminación: 36, 37, 38, 40, 43, 49, 50, 52, 60 y 62.

l) Carrera de Música:

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 19:

Tabla 19. Dificultad y discriminación para la carrera de Música.

Ítems	Encabezado	Media	Desv. Estándar	Discriminación
P38		0.50	0.51	0.23
P39		0.86	0.35	0.15
P40		0.64	0.49	0.17
P42		0.86	0.35	0.23
P43		0.86	0.35	0.32
P44		0.97	0.17	0.42
P48		0.78	0.42	0.50
P51		0.78	0.42	0.16
P52		0.64	0.49	0.29
P53		0.97	0.17	0.08
P54		0.56	0.50	0.24
P55		0.94	0.23	0.43
P59		0.64	0.49	0.17

En este caso, se removieron los ítems 35, 36, 37, 41, 45, 46, 47, 49, 50, 56, 57, 58, 60, 61, 62 y 63. La prueba final se construyó con los ítems restantes.

m) Carrera de Química:

Los resultados descriptivos y formales para la disciplina se disponen en la tabla 20:

Tabla 20. Dificultad y discriminación para la carrera de Química.

Ítems	Encabezado	Dificultad	Desv. Estándar	Discriminación
P37	Se solicita "Clasificar" todo lo que se	0.86	0.38	0.75
P38	John Dalton fue un científico que	0.86	0.38	0.32
P39	Dentro de un recipiente cerrado y	0.71	0.49	0.46
P41	Un cocinero hace el pan mezclando	0.43	0.53	0.24
P42	Considerando la Ley de las Proporciones	0.43	0.53	0.62
P43	La ley de Boyle-Mariotte, o ley de Boyle	0.86	0.38	0.75
P44	El ácido fosfórico (H ₃ PO ₄)	0.57	0.53	0.70
P45	El período de semidesintegración	0.43	0.53	0.62
P47	En el estudio de las reacciones químicas	0.71	0.49	0.88
P48	Observa la siguiente imagen:	0.14	0.38	0.45
P49	El círculo representa las moléculas	0.86	0.38	0.32

P51	Para conocer el impacto medioambiental	0.57	0.53	0.32
P53	Un estudiante realiza un experimento	0.71	0.49	0.26
P54	Un estudiante diseñó un experimento usando sal y agua	0.71	0.49	0.88
P56	El siguiente gráfico muestra la solubilidad	0.57	0.53	0.70
P57	Con base en el gráfico anterior	0.43	0.53	0.62
P60	En los climas fríos se requiere	0.71	0.49	0.26
P61	Las reacciones de neutralización	0.86	0.38	0.75
P62	La glucosa, uno de los componentes	0.86	0.38	0.75
P63	De acuerdo a los datos del ejercicio	0.71	0.49	0.88

Se removieron los ítems: 35, 36, 40, 46, 50, 52, 55, 58 y 59.

Cálculo de la fiabilidad por carrera.

Se obtuvo la fiabilidad mediante α de Cronbach para cada uno de los tests disciplinares pero con las pruebas reducidas con la composición que se ha informado previamente. Los resultados se disponen en la tabla 21.

Tabla 21. Fiabilidad de los tests (reducidos) de capacidades disciplinares de cada carrera.

Tests capacidades disciplinares	Fiabilidad
Biología	0.208
Educación Básica	0.657
Educación Diferencial	0.585
Educación Física	0.457
Educación Parvularia	0.437
Filosofía	0.594
Física	0.704
Historia	0.580
Inglés	0.669
Literatura y Ciencias Lenguaje	0.427
Matemáticas	0.698
Música	0.589
Química	0.918

Las informaciones siguen siendo, en la mayoría de los casos, más bajas que lo esperado, sin embargo, en todos los casos, la fiabilidad aumentó con la eliminación de los ítems problemáticos. Se pueden considerar en el borde de lo aceptable, aquellas pruebas que denotan una fiabilidad cercana a 0.70. Para mejorar a mejores niveles el desempeño de la fiabilidad, se requiere mejorar la estructura de construcción de los ítems.

Resultados finales por carrera.

Con los instrumentos reducidos, se han obtenido los resultados para cada una de las unidades académicas que participaron del estudio. En primer lugar se disponen los resultados de la prueba pedagógica por dimensión. Posteriormente los resultados de las pruebas disciplinares, para cerrar en definitiva con gráficos que resumen la información dispuesta.

De la prueba de capacidades pedagógicas.

Primero se informan tablas con las dimensiones del apartado actitudinal. Como toda escala de actitudes, no se informa en porcentajes, sino en puntuaciones, de 1 a 5 (Muy en Desacuerdo a Muy de Acuerdo, respectivamente). De modo que las medias y todos los resultados obedecen a esta escala.

Tabla 22. Dimensión disposición a trabajar con niños de la escala actitudinal.

Escala actitudinal prueba pedagógica					
Dimensión disposición a trabajar con niños					
Carreras	n	Media	DE	Mínimo	Máximo
Biología	31	4.3	0.4	3.6	5.0
Educación Básica	43	4.6	0.3	3.8	5.0
Educación Diferencial	58	4.4	0.3	3.6	5.0
Educación Física	85	4.3	0.4	3.0	5.0
Educación Parvularia	39	4.5	0.3	3.8	5.0
Filosofía	9	4.2	0.5	3.2	4.8
Física	41	4.0	0.4	3.2	4.8
Historia	71	4.2	0.5	2.4	5.0
Inglés	80	4.3	0.4	3.0	5.0
Ciencias del Lenguaje	63	4.4	0.3	3.6	5.0
Matemática	60	4.3	0.4	3.2	5.0
Química	7	4.3	0.3	3.8	4.8
Total	587	4.3	0.4	2.4	5.0

Los valores tienden a ser bastante altos, en particular entre las carreras propias de pedagogía. Algo más bajos en carreras asociadas con ciencias básicas.

Respecto a la dimensión respeto y tolerancia a la diversidad, los resultados se disponen en la tabla 23.

Tabla 23. Dimensión respeto y tolerancia a la diversidad de la escala actitudinal.

Escala actitudinal prueba pedagógica					
Carreras	Dimensión respeto y tolerancia a la diversidad				
	n	Media	DE	Mínimo	Máximo
Biología	31	4.3	0.6	2.8	5.0
Educación Básica	43	4.7	0.3	3.4	5.0
Educación Diferencial	58	4.8	0.3	3.4	5.0
Educación Física	85	4.5	0.4	3.4	5.0
Educación Parvularia	39	4.8	0.3	3.6	5.0
Filosofía	9	4.2	0.8	2.4	5.0
Física	41	4.2	0.5	2.4	5.0
Historia	71	4.4	0.5	3.2	5.0
Inglés	80	4.5	0.4	3.2	5.0
Ciencias del Lenguaje	63	4.6	0.4	3.8	5.0
Matemática	60	4.4	0.4	3.2	5.0
Química	7	4.4	0.4	3.8	5.0
Total	587	4.5	0.5	2.4	5.0

Los resultados obtenidos en esta dimensión son también elevados, en particular en carreras de pedagogía propiamente tales, pero más bajos en carreras como Biología, Filosofía y Física.

Tabla 24. Dimensión empatía de la escala actitudinal.

Escala actitudinal prueba pedagógica					
Carreras	Dimensión empatía				
	n	Media	DE	Mínimo	Máximo
Biología	31	3.8	0.6	2.2	4.8
Educación Básica	43	3.8	0.5	2.8	4.6
Educación Diferencial	58	3.9	0.5	2.4	5.0
Educación Física	85	3.6	0.6	2.4	4.8
Educación Parvularia	39	3.9	0.6	1.8	4.6
Filosofía	9	3.7	0.7	2.6	4.8
Física	41	3.7	0.6	2.4	4.8
Historia	71	3.5	0.7	1.0	4.8
Inglés	80	3.9	0.6	2.6	5.0
Ciencias del Lenguaje	63	3.7	0.6	2.6	5.0
Matemática	60	3.7	0.7	2.4	5.0
Química	7	3.3	0.9	1.8	4.2
Total	587	3.7	0.6	1.0	5.0

Los resultados de la dimensión empatía son, comparativamente, los más bajos de las tres dimensiones consultadas, dentro de esa variabilidad, los resultados obtenidos se muestran en la tabla 24.

Los resultados son más elevados en carreras de pedagogía propiamente tales, aunque ninguna supera los cuatro puntos. En el grado más bajo se ubica Historia y Química.

Tabla 25. Puntuaciones totales de la escala actitudinal.

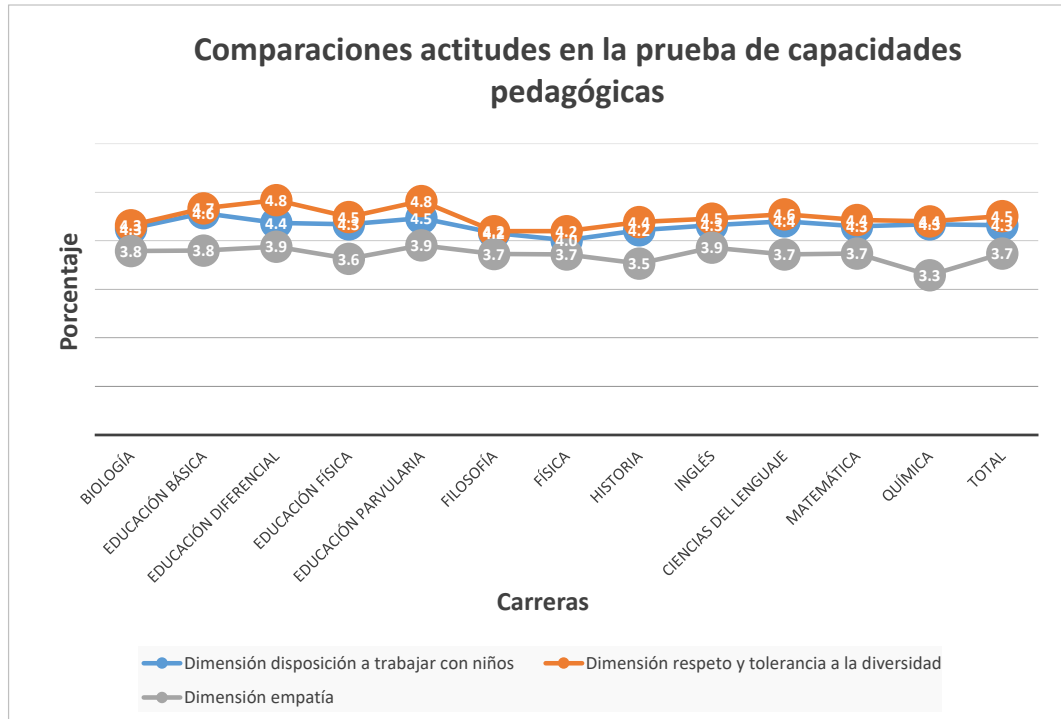
Escala actitudinal prueba pedagógica					
Carreras	n	Total actitudes pedagógicas			
		Media	DE	Mínimo	Máximo
Biología	31	4.4	0.4	3.1	5.0
Educación Básica	43	4.7	0.3	3.8	5.1
Educación Diferencial	58	4.7	0.3	3.9	5.2
Educación Física	85	4.4	0.3	3.8	5.0
Educación Parvularia	39	4.7	0.3	4.1	5.1
Filosofía	9	4.3	0.6	3.4	4.9
Física	41	4.2	0.4	2.9	4.9
Historia	71	4.3	0.4	3.1	5.1
Inglés	80	4.5	0.3	3.8	5.2
Ciencias del Lenguaje	63	4.5	0.3	3.9	5.1
Matemática	60	4.4	0.4	3.7	5.1
Química	7	4.3	0.4	3.8	4.8
Total	587	4.5	0.4	2.9	5.2

La última tabla de este apartado es del total de la escala de actitudes y al final se dispone un gráfico comparativo entre dimensiones.

Los resultados entre carreras no varían sustancialmente al considerar la puntuación total y las diferencias observadas son consecuencia de las diferencias previamente informadas. Destaca en particular Educación Básica y Diferencial por sus puntuaciones elevadas y Física por estar en el otro extremo, hacia puntuaciones bajas.

El gráfico de la figura 1 dispone una comparativa de los resultados en las tres dimensiones anteriormente dispuestas.

Figura 1. Comparación de dimensiones de la escala actitudinal de la prueba pedagógica.



Como se aprecia, y estando todas las dimensiones en la misma escala, la de mayor relevancia por sus puntuaciones comparativamente más altas, es la dimensión respeto y tolerancia a la diversidad. En esta dimensión destaca Educación Diferencial con el resultado más alto de todas las carreras. La dimensión más baja es empatía y de las carreras, Química está en el nivel más bajo.

A continuación se disponen las dimensiones relativas al apartado de desempeño en las capacidades transversales medidas por la prueba pedagógica. Los resultados se informan en porcentajes. Se disponen las primeras cuatro tablas con las dimensiones, la última informa de los resultados totales en esta parte de la prueba. Luego se presenta un gráfico comparativo entre dimensiones.

Tabla 26. Dimensión reflexión crítica del apartado capacidades pedagógicas.

Prueba de capacidades pedagógicas					
Carreras	n	Reflexión Crítica			
		Media	DE	Mínimo	Máximo
Biología	31	90.32	15.38	60	100
Educación Básica	43	89.77	15.35	40	100
Educación Diferencial	58	88.97	14.59	40	100
Educación Física	85	88.71	19.87	0	100
Educación Parvularia	39	92.31	14.23	40	100
Filosofía	9	93.33	14.14	60	100
Física	41	94.63	10.02	60	100
Historia	71	91.27	14.24	40	100
Inglés	80	89	16.5	40	100
Ciencias del lenguaje	63	89.84	15.61	40	100
Matemáticas	60	94	11.23	60	100
Música	36	93.89	14.2	40	100
Química	7	91.43	10.69	80	100
Total	623	90.85	15.23	0	100

Al igual que en otros resultados, aquí también se aprecia un resultado similar entre las distintas carreras. Sin embargo, la magnitud del logro en capacidad de reflexión crítica es mayor en la carrera de Física, seguida de Música. En el otro extremo está Educación Diferencial y Educación Física, aun cuando las diferencias son pequeñas.

Respecto de la dimensión Competencias Digitales, los resultados se disponen en la tabla 27.

Tabla 27. Dimensión competencias digitales del apartado capacidades pedagógicas.

Prueba de capacidades pedagógicas					
Carreras	n	Competencias digitales			
		Media	DE	Mínimo	Máximo
Biología	31	73.12	20.04	33.3	100.0
Educación Básica	43	75.19	16.81	33.3	100.0
Educación Diferencial	58	77.01	15.24	33.3	100.0
Educación Física	85	67.84	19.72	16.7	100.0
Educación Parvularia	39	74.79	14.24	33.3	100.0
Filosofía	9	68.52	24.22	33.3	100.0
Física	41	61.79	23.64	16.7	100.0
Historia	71	64.08	17.96	16.7	100.0
Inglés	80	71.25	18.75	0.0	100.0

Ciencias del lenguaje	63	67.99	17.79	16.7	100.0
Matemáticas	60	63.33	20.77	16.7	100.0
Música	36	60.65	18.32	16.7	83.3
Química	7	69.05	17.82	33.3	83.3
Total	623	68.7	19.21	0.0	100.0

Los resultados en esta dimensión son claramente más bajos que en la dimensión anterior. Quienes denotan mayores competencias digitales son los estudiantes de carreras propiamente pedagógicas, mientras que los porcentajes más bajos están entre los estudiantes de Física y Música.

Los resultados de la dimensión Metacognición en comprensión de textos se disponen en la tabla 28.

Tabla 28. Dimensión metacognición con comprensión de textos del apartado capacidades pedagógicas.

Prueba de capacidades pedagógicas					
Carreras	n	Metacognición en comprensión de textos			
		Media	DE	Mínimo	Máximo
Biología	31	35.48	19.12	0	67
Educación Básica	43	44.96	20.42	0	67
Educación Diferencial	58	51.15	23.54	0	100
Educación Física	85	40.78	23.78	0	100
Educación Parvularia	39	43.59	18.97	0	67
Filosofía	9	48.15	29.4	0	100
Física	41	44.72	23.11	0	100
Historia	71	47.89	25.65	0	100
Inglés	80	43.75	24.07	0	100
Ciencias del lenguaje	63	51.85	21.4	0	100
Matemáticas	60	42.78	22.21	0	100
Música	36	41.67	23.06	0	100
Química	7	52.38	26.23	0	67
Total	623	45	23.14	0	100

Se evidencian resultados sustancialmente más bajos que en las anteriores dimensiones. Entre las carreras de mejor desempeño está Educación Diferencial, Ciencias del Lenguaje y Química. Entre los más bajos se ubica Biología, con un resultado que no sobrepasa el 36% de la prueba correcta.

Los resultados para la dimensión Comprensión de Textos Escritos se dispone en la tabla 29.

Tabla 29. Dimensión comprensión de textos escritos del apartado capacidades pedagógicas.

Prueba de capacidades pedagógicas					
Comprensión de textos escritos					
Carreras	n	Media	DE	Mínimo	Máximo
Biología	31	40.65	20.32	0	80
Educación Básica	43	50.7	20.17	20	80
Educación Diferencial	58	54.83	20.02	20	100
Educación Física	85	41.41	24.26	0	100
Educación Parvularia	39	45.13	18.76	0	80
Filosofía	9	42.22	25.39	0	80
Física	41	34.15	16.28	0	60
Historia	71	45.63	20.34	0	100
Inglés	80	40.75	20.97	0	100
Ciencias del lenguaje	63	40	16.85	0	80
Matemáticas	60	46.67	15.91	0	80
Música	36	41.67	23.11	0	80
Química	7	42.86	17.99	20	60
Total	623	43.82	20.61	0	100

Esta dimensión también denota un bajo desempeño general. El mejor resultado es para la carrera de Educación Diferencial. Entre los resultados más bajos está Física y, curiosamente, Ciencias del Lenguaje.

Los resultados totales para la dimensión capacidades pedagógicas se disponen en la tabla 30.

Tabla 30. Porcentajes totales de logro en el apartado capacidades pedagógicas de la prueba pedagógica.

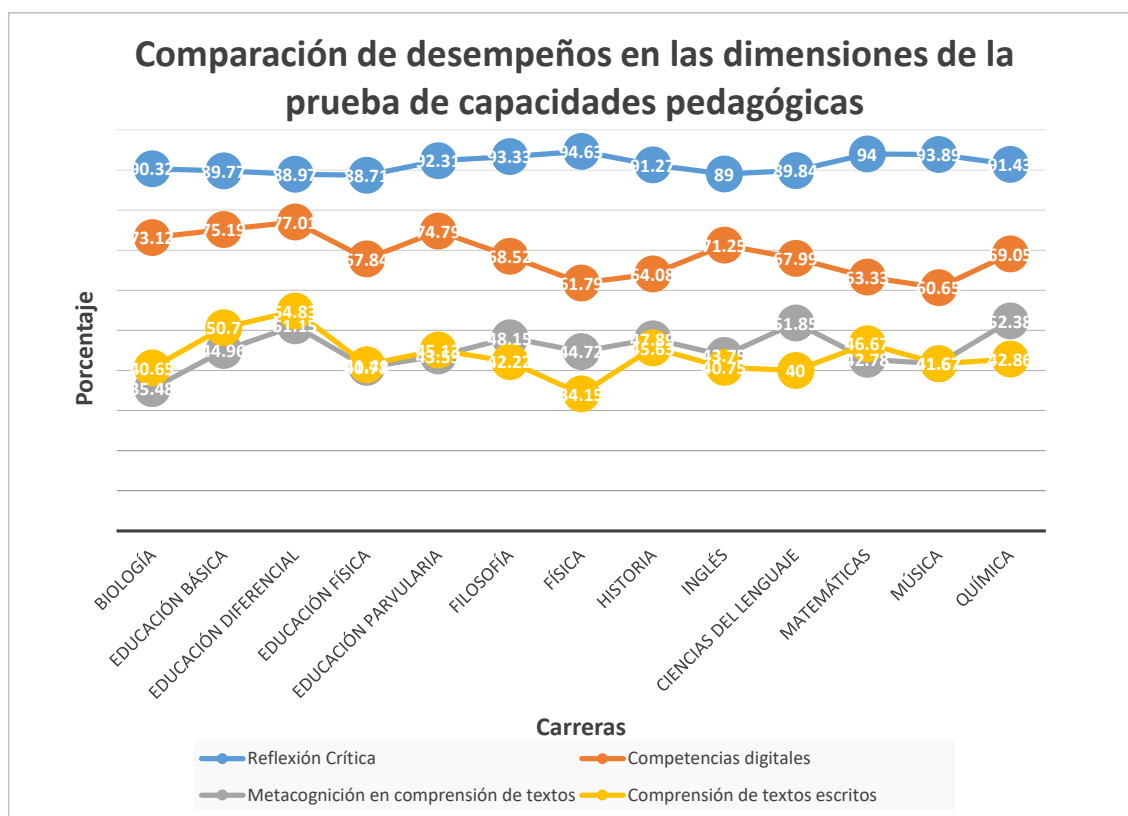
Prueba de capacidades pedagógicas					
Total capacidades pedagógicas					
Carreras	n	Media	DE	Mínimo	Máximo
Biología	31	63.2	10.4	31.6	84.2
Educación Básica	43	67.8	9.1	47.4	89.5
Educación Diferencial	58	70.2	10.7	47.4	89.5
Educación Física	85	62.1	12.9	15.8	89.5
Educación Parvularia	39	66.7	8.2	52.6	84.2
Filosofía	9	64.9	7.4	52.6	78.9
Física	41	60.5	10.0	36.8	78.9
Historia	71	63.8	9.9	31.6	89.5
Inglés	80	63.6	11.5	15.8	89.5
Ciencias del lenguaje	63	63.8	10.3	31.6	84.2

Matemáticas	60	63.8	8.3	42.1	84.2
Música	36	61.4	10.8	36.8	84.2
Química	7	65.4	12.8	42.1	78.9
Total	623	64.2	10.7	15.8	89.5

El mejor desempeño lo tiene Educación Diferencial, con un 70% de la prueba lograda considerando todas las dimensiones. Lo sigue Educación Básica y posteriormente Educación Parvularia. En el otro extremo del desempeño, están Música y Biología.

La figura 2 dispone una comparación de porcentajes por cada una de las dimensiones de acuerdo a las carreras evaluadas.

Figura 2. Comparación de dimensiones de capacidades pedagógicas de la prueba pedagógica.



Hay una clara diferencia en los desempeños que implican comprensión de textos y metacognición, y la capacidad de reflexión crítica y las competencias digitales. Los estudiantes en general, denotan buen nivel en su capacidad de reflexión, algo menos en sus capacidades digitales, aunque en general dentro de lo esperado (60 o 65%) y claramente por debajo de lo necesario en comprensión de textos.

De la prueba de capacidades disciplinares.

La evaluación específica de cada carrera considera un número variable de ítems por prueba. Las dimensiones que las componen son conocidas y definidas por cada unidad académica. Aquí se informa del resultado total habiendo eliminado todos los ítems con dificultades en su comportamiento psicométrico. Los resultados se disponen en la tabla 31.

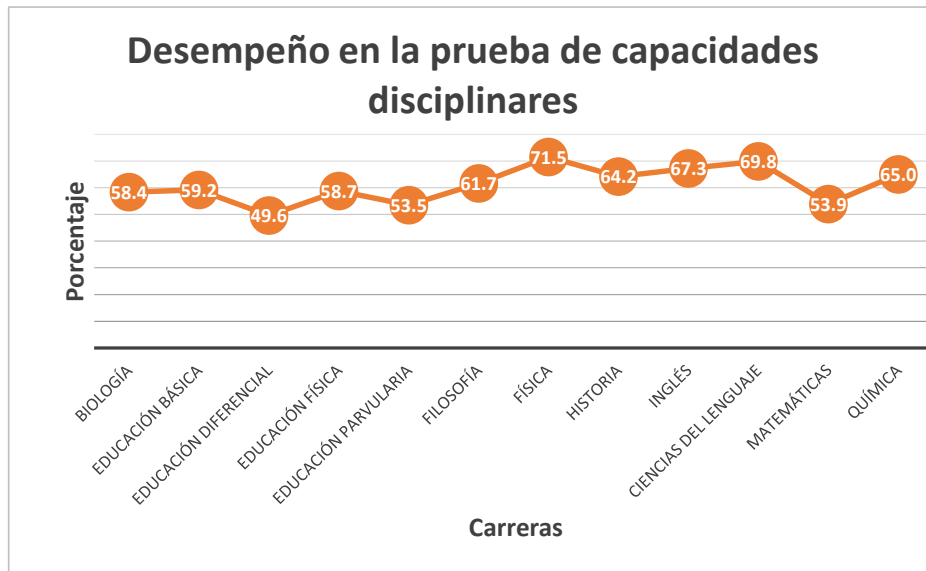
Tabla 30. Porcentajes totales de logro en la prueba de capacidades disciplinares.

Prueba de capacidades disciplinares					
Carreras	n	Media	DE	Mínimo	Máximo
Biología	31	58.4	11.6	36.8	78.9
Educación Básica	43	59.2	18.6	20.0	93.3
Educación Diferencial	58	49.6	15.1	11.1	77.8
Educación Física	85	58.7	14.1	21.4	92.9
Educación Parvularia	39	53.5	14.9	21.4	85.7
Filosofía	9	61.7	15.0	35.0	85.0
Física	41	71.5	18.4	12.5	100.0
Historia	71	64.2	14.7	0.0	100.0
Inglés	80	67.3	16.0	9.1	95.5
Ciencias del Lenguaje	63	69.8	17.1	30.0	100.0
Matemáticas	60	53.9	17.0	15.0	95.0
Química	7	65.0	29.4	10.0	100.0
Total	587	61.0	17.3	0.0	100.0

Los resultados están expresados en porcentajes de logro. En este nivel el desempeño de las carreras es bien distinto al obtenido en la prueba de capacidades pedagógicas. Aquí, los mejores resultados son obtenidos por Física, Ciencias del Lenguaje, seguido de Inglés. Los resultados más bajos están entre Educación Diferencial, Educación Parvularia y Matemáticas.

La figura 3 dispone el resultado de la tabla anterior en un gráfico resumen.

Figura 3. Resultados finales de la prueba de capacidades disciplinares.



El gráfico sólo corrobora lo que se ha expuesto anteriormente en la tabla. En términos generales son pocas las carreras con desempeños mayores a 60%.

Síntesis.

El presente informe tuvo por objetivo disponer dos informaciones de relevancia: a) La composición definitiva de las distintas pruebas construidas, en función de los ítems que cumplen con las propiedades psicométricas mínimas para ser incorporados como parte de un instrumento de medición, b) los resultados finales de las distintas carreras en la prueba pedagógica, tanto en capacidades como en actitudes, y la prueba de capacidades disciplinares.

En cuanto a la composición, muchos de los tests debieron ser reducidos hasta en un 40%, removiendo los ítems inadecuados. El principal criterio fue la capacidad de discriminación del ítem, ya que valores menores a 0.3 no son aceptables. Sin embargo, dado que la eliminación bajo este criterio hubiese descartado una porción aún mayor de ítems, se tomó en cuenta también la dificultad del ítem. Con ambos criterios, se decidió la eliminación o conservación de los ítems.

Queda aún el análisis de estructura interna de los tests (validez de constructo), para lo que se necesita una muestra mayor que las actualmente disponibles por cada una de las carreras. No obstante, el análisis actual es una referencia mínima para construir ulteriormente, ítems que tengan una dificultad y una discriminación aceptable.

La fiabilidad calculada, como se puede observar en la tabla 21, salvo en las pruebas de Química, Física, Inglés y Matemáticas, requieren una cuidadosa revisión e incluso nuevas versiones, como ocurre con la prueba de Biología.

En cuanto a los resultados finales, en la prueba de capacidades pedagógicas, los mejores resultados están entre las carreras propiamente pedagógicas como Educación Diferencial, Básica y Parvularia. En el ámbito de la prueba de capacidades disciplinares, los desempeños más altos están en carreras de ciencias básicas como Física o Química. Cabe notar además, que son las pruebas con mejores propiedades psicométricas, por lo que las medidas obtenidas con ellos son al menos fiables.

Bibliografía

- Abad, F. J., Olea, J. y Ponsoda, V. (2001). Analysis of optimum number alternatives from the item response theory. *Psicothema*, 13(1), 152-158.
- Abad, F., Olea, J., Ponsoda, V. y García, C. (2010). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis.
- Ihaka, R. & Gentleman, R. (1996) R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299-314
- Martínez, M. R., Hernández, M. J. y Hernández, M. V. (2006). *Psicometría*. Madrid: Alianza Editorial.
- R Core Team (2017). *R: A language and environment for statistical computing*. Viena: R Foundation for Statistical Computing.
- Santisteban, C. (1990). *Psicometría. Teoría y práctica en la construcción de tests*. Madrid: Norma.