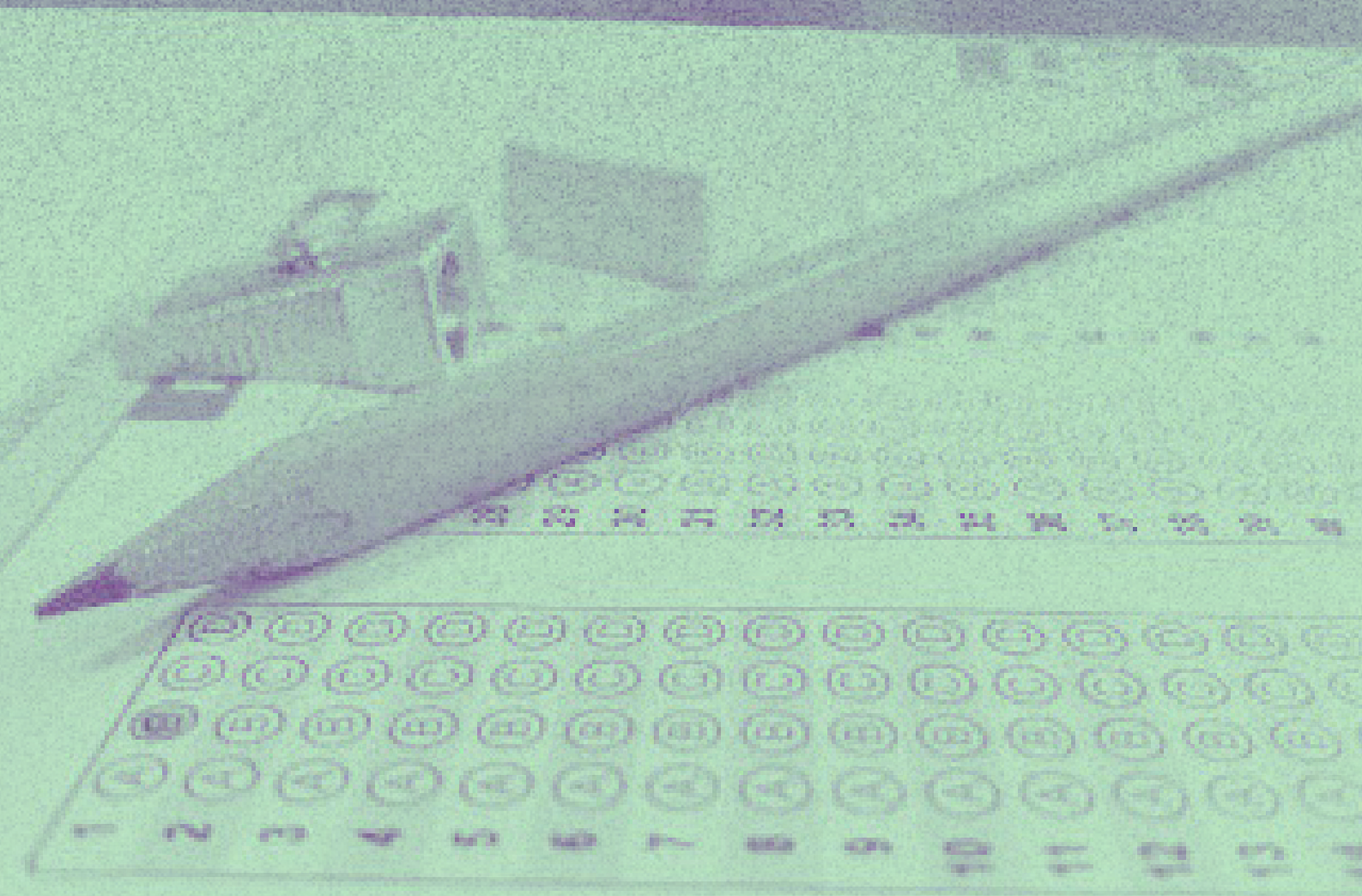
DICIEMBRE 2023

PERCEPCIÓN DEL IMPACTO DE LA UNIDAD DE EVALUACIÓN

SOBRE LA COMUNIDAD EDUCATIVA DE LA ESCUELA DE MEDICINA DE LA UNIVERSIDAD DE CHILE



FACULTAD DE MEDICINA
UNIVERSIDAD DE CHILE

AUTORES

NICOLÁS LAVADOS
EMILIA WINKLER
JOAQUIN LÓPEZ
FELIPE MONTENEGRO
CAMILO GÓMEZ
VALERIA SIERRA
OSCAR JEREZ

DOCENTES

ALVARO HERRERA
OSCAR JEREZ

DISEÑO

JOSEFA ALFARO

ÍNDICE

04 INTRODUCCIÓN

07 OBJETIVOS

08 METODOLOGÍA

13 RESULTADOS
DE ENCUESTA A ALUMNOS

22 RESULTADOS
DE ENTREVISTA A DOCENTES

23 BENCHMARKING
CASOS RELEVANTES EN OTRAS INSTITUCIONES EDUCATIVAS

25 DISCUSIÓN

27 RECOMENDACIONES

28 CONCLUSIONES

29 BIBLIOGRAFÍA

30 ANEXOS

INTRODUCCIÓN

lEl Observatorio de Educación Médica de la Universidad de Chile es un proyecto que nace con el objetivo de innovar en docencia y mantener el constante perfeccionamiento de los métodos de enseñanza en la Escuela de Medicina de la Universidad.En este sentido, la presente investigación pretende realizar una revisión general sobre la estandarización de las pruebas de opción múltiple ("multiple choice"), apuntando a la situación local en la Facultad, para luego elaborar recomendaciones atingentes que permitan mejorar la docencia.

Las evaluaciones tienen un rol fundamental dentro de la educación, ya que además de ser una herramienta de aprendizaje, permiten valorar la adquisición de ciertas habilidades y conocimientos en el estudiante (Sandoval et al., 2022). Es por ello que este boletín del Observatorio de Educación Médica se enfocará en ellas.

Las evaluaciones de opción múltiple son utilizadas a nivel internacional como método de selección y evaluación (Patterson et al., 2015). Dentro de sus virtudes, destaca que permiten evaluar de manera simultánea múltiples áreas de contenido y que se administran en cortos periodos de tiempo, lo que hace más sencilla su aplicación y estandarización (Epstein, 2007). No obstante, también tiene desventajas, por ejemplo, otros tipo de evaluación permiten una mayor elaboración de la respuesta, la escritura de los enunciados implica más dificultad y tiempo, puede inducir la respuesta correcta a través de las otras alternativas ofrecidas, entre otros (Epstein, 2007).

Las "pruebas estandarizadas" se denominan así a propósito de que buscan generar un estándar sistematizado tanto en la creación y aplicación de la prueba como en la interpretación que se hace de los datos obtenidos, de modo que un mismo instrumento, aplicado de igual manera a los estudiantes, permita obtener información válida y comparable entre ellos, incluso respecto a otros cursos, instituciones y/o países (SEPAUC, 2017). La estandarización de evaluaciones de opción múltiple no sólo permite disminuir la subjetividad de las valoraciones, sino que además tiene un rol en la determinación de tendencias en los resultados de los alumnos y permite evaluar el desempeño de los profesores (Hernández et al., 2018). De esta manera, las evaluaciones de opción múltiple estandarizadas son una herramienta para establecer metas educativas, diseñar planes y programas de estudio, prevenir deficiencias de aprendizaje en cursos posteriores y tomar decisiones relacionadas con las políticas internas de evaluación (Hernández et al., 2018).

El proceso de estandarización presenta una gran dificultad bajo esta definición, requiriendo tiempos prolongados para el diseño de cada prueba. Sin embargo, en la Facultad de Medicina de la Universidad de Chile, este se lleva a cabo mediante el uso de protocolos y procedimientos para que

las evaluaciones sean justas y comparables entre sí, sin necesidad de ceñirse a las guías de estandarización que permiten la comparación entre países, de manera que sea posible elaborar las evaluaciones en los tiempos adecuados para el ritmo de la universidad.

En la Facultad de Medicina de la Universidad de Chile, la estandarización de pruebas de opción múltiple es llevada a cabo por medio de la Unidad de Evaluación (UNEV). Esta entidad, creada en el año 2000, despliega protocolos que agilizan el proceso de elaboración y validación de evaluaciones, con el propósito de enriquecer la manera en que se evalúan los conocimientos y habilidades pertinentes para cada asignatura.

Este protocolo consta de cuatro fases esenciales. En la primera, se aborda la confección de preguntas. Aquí se generan preguntas acordes al programa de estudios de cada asignatura, diseñando al menos una pregunta para cada tipo de contenido a evaluar. En este proceso se incluyen temas que se encuentran en el plan de estudio, mientras que se excluyen aquellos ajenos al programa, evitando su evaluación. Además, se priorizan las temáticas según su relevancia.

La segunda fase se concentra en garantizar la calidad técnica del proceso. Se cuenta con personal debidamente formado para confeccionar las pruebas y supervisar su integridad técnica.

La validación de contenidos constituye la tercera etapa, en la cual se someten las preguntas a revisión. Al menos dos Profesores Encargados de Curso (PECs) colaboran en este proceso, quienes identifican posibles errores de contenido y realizan ajustes en caso de ser necesario, asegurando así la calidad y pertinencia de las preguntas.



El último paso se refiere a la aplicación y retroalimentación del examen. En esta instancia, se desarrollan dos versiones del examen para cada prueba, garantizando uniformidad en ambas versiones para todos los estudiantes. Además, en esta etapa se provee retroalimentación efectiva por parte del Equipo Docente.

Es importante señalar que la UNEV establece un límite máximo de 48 horas para la revisión. Este plazo se inicia en el momento en que las hojas de evaluación son recibidas por la Escuela de Medicina.

Sin embargo, a pesar de que la UNEV se encuentra en funcionamiento desde el año 2000, son pocas las asignaturas que validan sus evaluaciones con esta unidad. Este hecho resalta la necesidad de comparar el desempeño y las impresiones de los estudiantes entre pruebas estandarizadas y no estandarizadas. El propósito de esta comparación es destacar las diferencias existentes entre estos dos enfoques.

Aunque la literatura presenta resultados diversos en relación con la importancia de la percepción de los estudiantes sobre la calidad de su aprendizaje, se subraya el impacto significativo que esta variable tiene en su nivel de involucramiento con los estudios. Esta percepción puede influir en si abordan el aprendizaje de manera superficial o profunda, lo que, a su vez, afecta directamente el desarrollo de las competencias esperadas. Además, es relevante destacar que los métodos utilizados para evaluar esta percepción pueden cambiar la orientación de los estudiantes hacia un enfoque de estudio más riguroso (Villa et al., 2020).

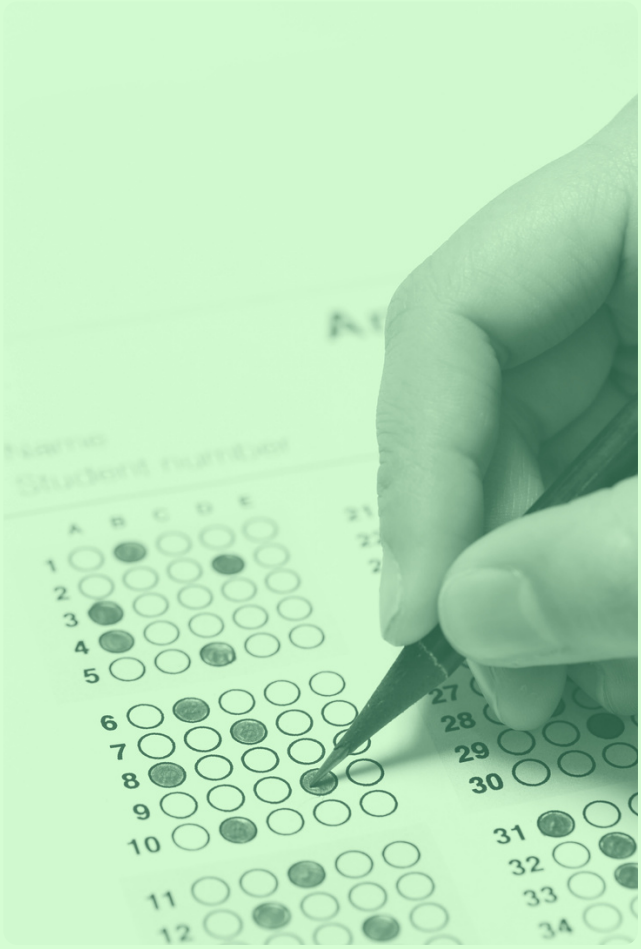
En este contexto, el presente boletín busca analizar el impacto de la UNEV sobre la calidad de las evaluaciones de selección múltiple, tanto para los estudiantes como para los profesores.



OBJETIVOS

OBJETIVO GENERAL

- Analizar el impacto de la Unidad de Evaluación (UNEV) sobre la calidad de las evaluaciones de selección múltiple en diversas asignaturas de la carrera de Medicina en la Facultad de Medicina de la Universidad de Chile.



OBJETIVOS ESPECÍFICOS

- Contrastar la percepción de los estudiantes en asignaturas que validan sus evaluaciones de selección múltiple con la UNEV en comparación con aquellas que no lo hacen en segundo, tercer y quinto año de medicina.
- Conocer la percepción de los profesores encargados de curso (PEC) que participan en la UNEV.
- Difundir la labor de la UNEV entre el cuerpo docente de la carrera, promoviendo su conocimiento y entendimiento.
- Analizar las opiniones y experiencias de los estudiantes con respecto a la metodología evaluativa de las asignaturas elegidas en este estudio.
- Generar recomendaciones para la comunidad educativa de la Facultad de Medicina con respecto al uso de la UNEV.

DISEÑO DE ESTUDIO

Se llevó a cabo un estudio transversal con el objetivo de conocer las percepciones de los estudiantes de Medicina de la Universidad de Chile con respecto a las siguientes asignaturas cuya metodología evaluativa se ciñe al protocolo de la Unidad de Evaluación (UNEV) de la Escuela de Medicina de la Universidad de Chile, Semiología I y II, Medicina Interna I y II y Ginecología y Obstetricia, y contrastar estas opiniones con respecto a aquellas que no se encuentran en la UNEV: Especialidades Médico-Quirúrgicas II, Fisiopatología I y II, Fisiología I y II y Farmacología I y II.

POBLACIÓN Y MUESTRA

La población objetivo consistió en estudiantes de Medicina de la Universidad de Chile. Inicialmente, se identificó una población de 1752 estudiantes, cifra obtenida a partir de los registros de votaciones de UCampus. Para evitar sesgos, se decidió excluir los ingresos de matrículas de primer y segundo año (256 y 261 estudiantes, respectivamente) debido a que, a la fecha del estudio, los estudiantes de estas cohortes no han finalizado asignaturas cuyas evaluaciones se encuentren validadas por la Unidad de Evaluación. Además se restaron las cohortes de sexto y séptimo año, debido al nulo interés de estos niveles a participar en el estudio. Por lo tanto, la población fue de 709 estudiantes.

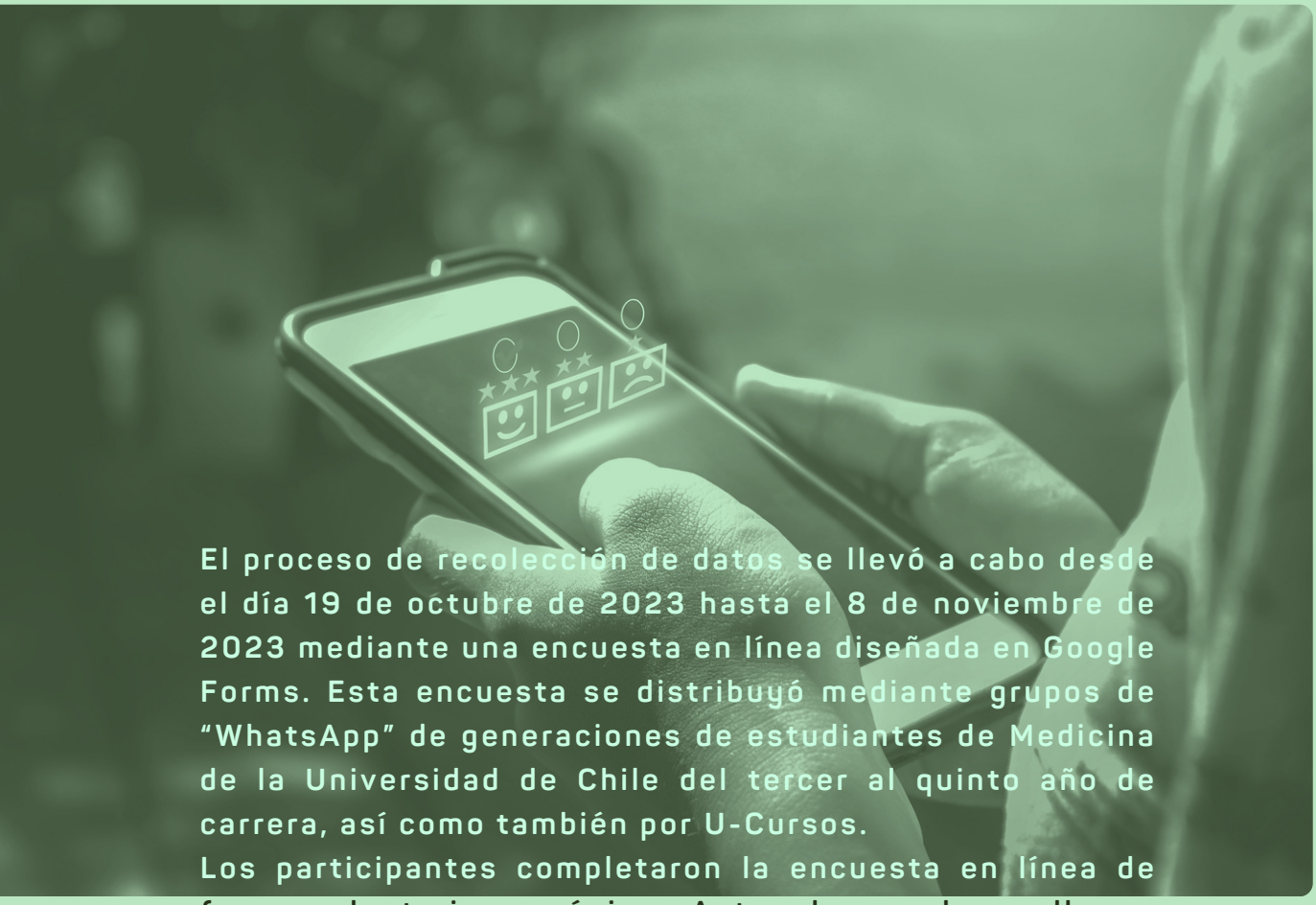
Para el cálculo del tamaño de muestra, se utilizó un nivel de confianza del 95%, y un margen de error del 5%.

Así, se determinó que, para obtener resultados confiables, se requiere una muestra mínima de 249 estudiantes. Esto se calculó utilizando la fórmula del tamaño de muestra para muestreo estratificado en poblaciones finitas:

$$n = \frac{NZ^2S^2}{d^2(N-1)+Z^2S^2}$$

Donde: n = tamaño de la muestra
N = tamaño de la población, Z = valor de Z crítico, S^2 = varianza de la población en estudio, d = nivel de precisión absoluta, referido a la amplitud del intervalo de confianza deseado en la determinación del valor promedio de la variable en estudio.

RECOLECCIÓN DE DATOS



El proceso de recolección de datos se llevó a cabo desde el día 19 de octubre de 2023 hasta el 8 de noviembre de 2023 mediante una encuesta en línea diseñada en Google Forms. Esta encuesta se distribuyó mediante grupos de "WhatsApp" de generaciones de estudiantes de Medicina de la Universidad de Chile del tercer al quinto año de carrera, así como también por U-Cursos.

Los participantes completaron la encuesta en línea de forma voluntaria y anónima. Antes de acceder a ella, se les proporcionó una explicación detallada sobre los objetivos y el propósito del estudio. Los estudiantes que aceptaron participar en el estudio fueron invitados a completar el cuestionario escribiendo su correo y haciendo clic en el enlace proporcionado.

Además, se solicitó la opinión de los docentes que participan activamente en la UNEV sobre su experiencia con esta. El contacto se realizó mediante sus respectivos correos institucionales.

En la primera parte de la encuesta, se recopiló el año de estudio de los participantes. Esta variable de segmentación nos permitió analizar las respuestas en función de esta característica.

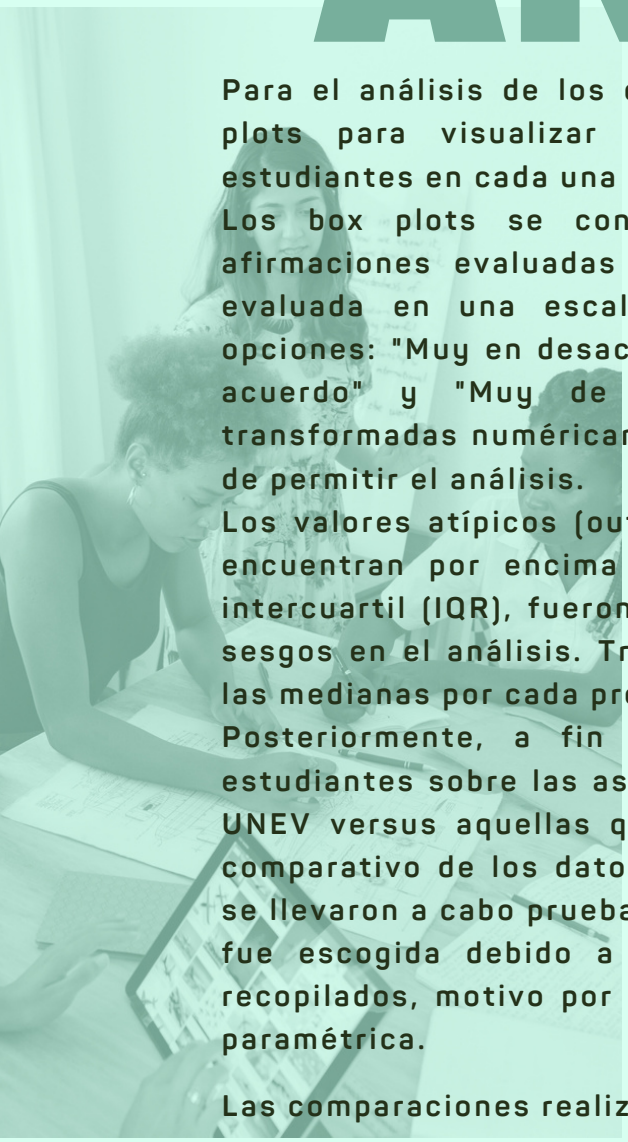
En la segunda parte se proporcionaron 7 afirmaciones, donde los participantes podían indicar si se encontraban "Muy en desacuerdo", "En desacuerdo", "Neutral", "De acuerdo" y "Muy de acuerdo", según su propia percepción. Las afirmaciones fueron:

Dimensión	Variable	Definición de la Variable	Ítems (Preguntas de la Encuesta)
Validez	Pertinencia del Contenido	Concordancia entre los ítems de la prueba y el currículo (Bloom, 1956; Wiggins, 1998)	1. Las preguntas de las evaluaciones eran acordes a los contenidos del curso.
Fiabilidad	Cantidad de Preguntas	Número apropiado de ítems en la prueba para medir habilidades (Nunnally, 1978)	2. La cantidad de preguntas de las evaluaciones era apropiada.
Eficiencia	Adecuación del Tiempo Asignado	Tiempo asignado para completar la prueba (Frederiksen, 1984)	3. La cantidad de tiempo asignado para responder las evaluaciones era apropiada.
Retroalimentación	Oportunidad y Adecuación de Retroalimentación	Calidad y tiempo de la retroalimentación recibida (Black & Wiliam, 1998; Hattie & Timperley, 2007)	4. Recibí retroalimentación adecuada y oportuna luego de las evaluaciones.
Interpretación	Correspondencia Nota-Aprendizaje	Grado en que la calificación refleja el nivel de aprendizaje (Sadler, 1989)	5. Siento que mis calificaciones en la asignatura reflejaron mi aprendizaje.
Claridad	Claridad en la Redacción	Facilidad para comprender los ítems de la prueba (Haladyna & Downing, 1989)	6. Las preguntas de las evaluaciones estaban redactadas de manera clara y no causaran confusión.
Metodología	Adecuación de la Metodología	Conveniencia de los métodos y criterios de evaluación (Biggs, 1996; Shepard, 2000)	7. La metodología evaluativa y requisitos de aprobación de la asignatura fueron apropiados.

Finalmente, se dejó un espacio para que los estudiantes comenten su apreciación con respecto a estas asignaturas al final de la encuesta.

DISEÑO DE LA ENCUESTA

ANÁLISIS DE DATOS



Para el análisis de los datos, se utilizó el método de los box plots para visualizar y comparar las respuestas de los estudiantes en cada una de las afirmaciones.

Los box plots se construyeron para cada una de las 7 afirmaciones evaluadas en la encuesta. Cada afirmación fue evaluada en una escala ordinal que incluía las siguientes opciones: "Muy en desacuerdo", "En desacuerdo", "Neutral", "De acuerdo" y "Muy de acuerdo". Estas respuestas fueron transformadas numéricamente del 1 al 5, respectivamente, a fin de permitir el análisis.

Los valores atípicos (outliers), definidos como aquellos que se encuentran por encima o por debajo de 1.5 veces el rango intercuartil (IQR), fueron identificados y eliminados para evitar sesgos en el análisis. Tras esto, se calcularon los promedios y las medianas por cada pregunta para cada asignatura.

Posteriormente, a fin de comparar las respuestas de los estudiantes sobre las asignaturas que siguen el protocolo de la UNEV versus aquellas que no lo hacen, se realizó un análisis comparativo de los datos obtenidos en la encuesta. Para esto, se llevaron a cabo pruebas U de Mann-Whitney en R. Esta prueba fue escogida debido a la falta de normalidad en los datos recopilados, motivo por el cual se debió utilizar una prueba no paramétrica.

Las comparaciones realizadas fueron:

En segundo año:
- Semiología I y II y Fisiología I y II

En tercer año:
- Medicina Interna I y II y Fisiopatología I y II
- Medicina Interna I y II y Farmacología I y II

En quinto año:
- Ginecología y Obstetricia y Especialidades Médico-Quirúrgicas II.

En el análisis, se calculó el p-value obtenido para cada pregunta, y los resultados de esta comparación se expresaron en tablas.

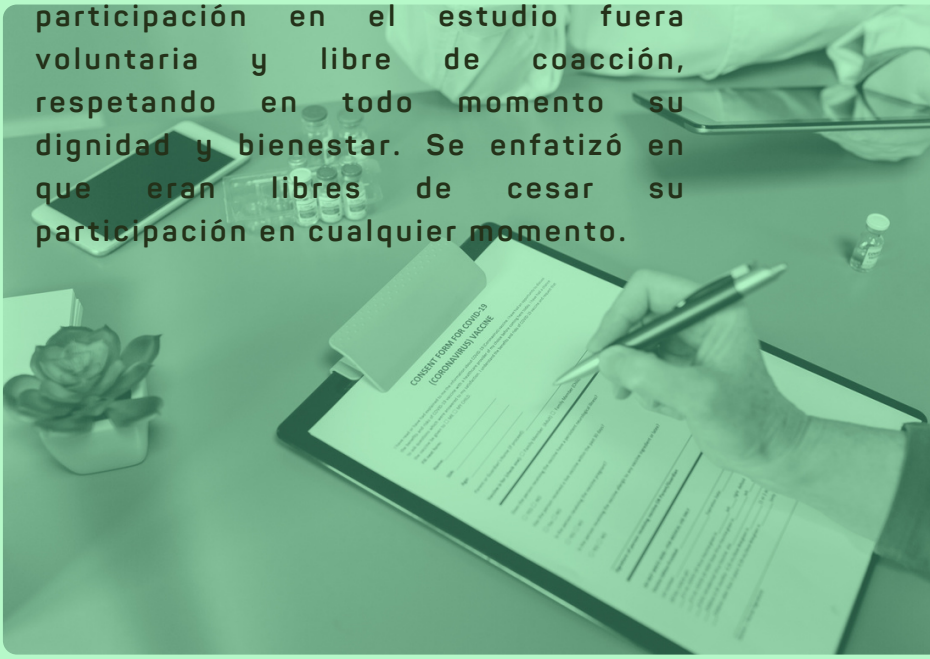
Además, se realizaron comparaciones de una misma asignatura según año de carrera, a fin de determinar si existen diferencias en sus percepciones.

CONSIDERACIONES ÉTICAS

Antes de llevar a cabo la recolección de datos, se implementó un proceso de consentimiento informado, tanto para docentes como para estudiantes, para así garantizar la ética y el respeto por los derechos de los participantes. Estos documentos se adjuntan en anexos.

No se recopilaron datos sensibles, tales como rut o dirección de los docentes participantes para asegurar su confidencialidad. Tampoco se recopiló nombre, rut, sexo ni dirección de los estudiantes que decidieron participar en el estudio. No se divulgará la identidad de ningún docente o estudiante en informes o publicaciones.

Tanto docentes como estudiantes fueron informados con respecto a sus derechos, y se aseguró de que su participación en el estudio fuera voluntaria y libre de coacción, respetando en todo momento su dignidad y bienestar. Se enfatizó en que eran libres de cesar su participación en cualquier momento.



RESULTADOS ENCUESTA ALUMNOS

En la encuesta realizada, participaron un total de 65 estudiantes de pregrado, de los cuales 20 se encontraban cursando tercer año, 19 cuarto año, y 16 quinto año al momento de responder la encuesta.

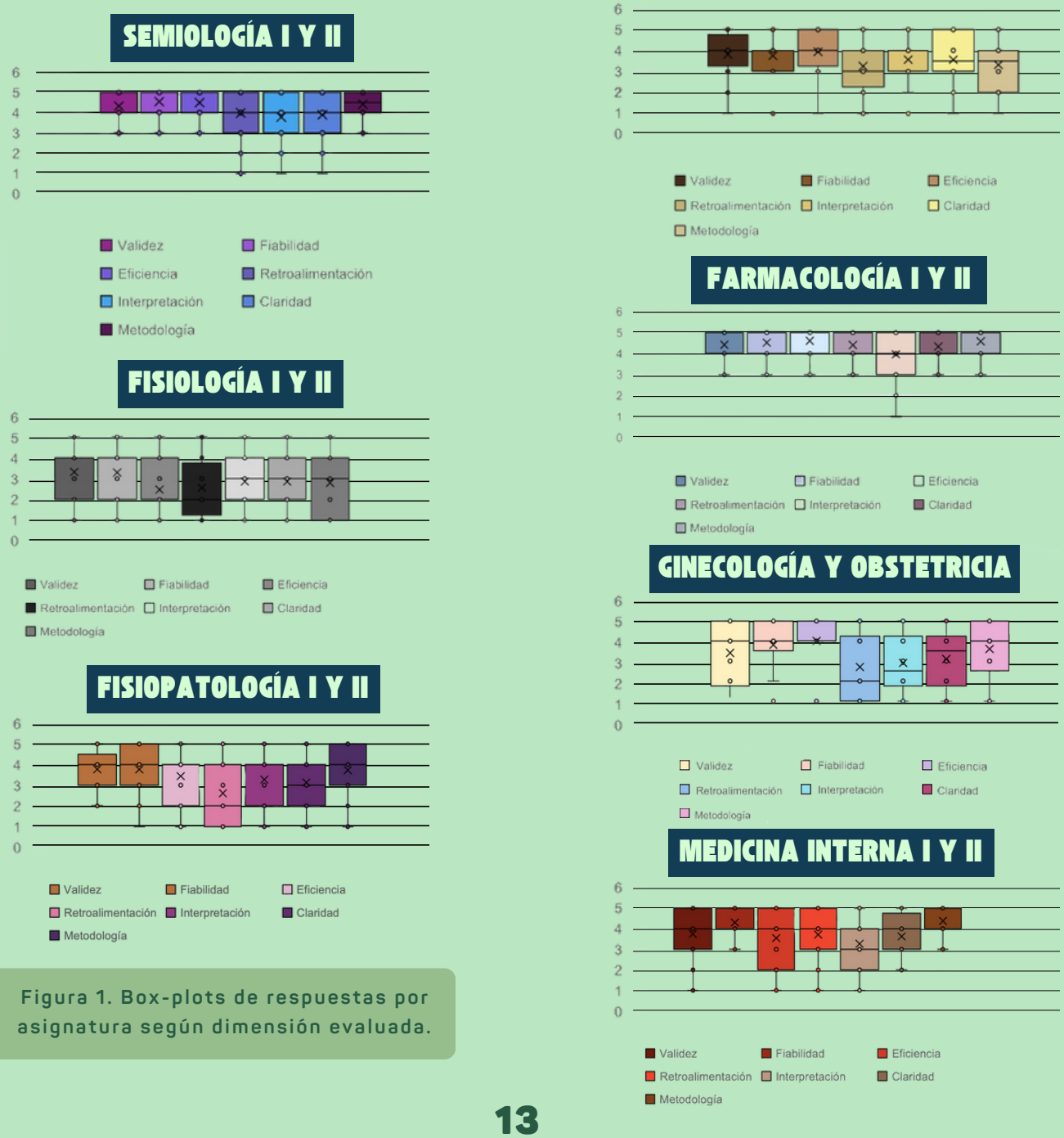


Figura 1. Box-plots de respuestas por asignatura según dimensión evaluada.

SEMIOLOGÍA

Dentro de las dimensiones evaluadas en Semiología I y II, se obtuvieron promedios que oscilan entre 3.78 y 4.55, correspondientes a interpretación y fiabilidad, respectivamente. Las medianas se mantienen entre 4 y 5.

FISIOLOGÍA

En las dimensiones evaluadas para esta asignatura, destacan promedios que van desde un 2.47, correspondiente a eficiencia, hasta un máximo de 3.32 en validez. Por su parte, las medianas se encuentran en un rango entre 2 y 4.

FISIOPATOLOGÍA

Los promedios obtenidos en Fisiopatología se encontraron entre 2.61 y 3.77, correspondientes a las dimensiones retroalimentación y validez, respectivamente. Las medianas, por su parte, se sitúan en el rango de 2 a 4.

MEDICINA INTERNA

En Medicina Interna, el promedio mínimo fue el de interpretación, con un valor de 3.27, y el máximo, de 4.37, correspondiente a metodología. Las medianas, por su parte, se mantuvieron entre 3 y 4 en todas las dimensiones.

FARMACOLOGÍA

Farmacología obtuvo promedios que van desde un 3.96 para la categoría de interpretación, hasta un máximo de 4.58 en metodología. Resaltan medianas entre 4 y 5 en todas las dimensiones analizadas.

GINECOLOGÍA Y OBSTETRICIA

Dentro de los promedios obtenidos en Ginecología y Obstetricia, destaca un mínimo de 3.1 en la dimensión claridad, y un máximo de 4.37 en fiabilidad. Las medianas se encontraban entre 2 y 4.

ESPECIALIDADES MÉDICO QUIRÚRGICAS II

Los promedios obtenidos en el análisis oscilan entre 3.25 para validez, y 3.93 para las dimensiones de fiabilidad y eficiencia. Las medianas se encontraron en un rango de 3 a 4.

COMPARATIVA ENTRE ASIGNATURAS DE SEGUNDO AÑO

En cuanto a los resultados de segundo año, se observa que el ramo Semiología I y II, asociado a la UNEV, se encuentra significativamente mejor evaluado en todas las áreas respecto a Fisiología I y II, ramo que no ciñe sus evaluaciones a los protocolos de la UNEV. Este último presenta valores promedio menores a 3 en cinco categorías, en contraste con Semiología, que presenta todos sus valores cercanos a 4 e incluso más. También se observa una mayor heterogeneidad en las respuestas en el primer caso. Si bien Fisiología se caracteriza por ser un ramo complejo y con mayor tasa de reprobación, los resultados permiten ver una clara diferencia entre ambos ramos, lo que permite plantear la hipótesis de que la UNEV sería de ayuda para mejorar la calidad y percepción de las evaluaciones. Es importante destacar que para ambos se tiene una cantidad de respuestas igual o muy cercana al valor total de la muestra (n=65).

	Semiología I y II (n = 65)	Fisiología I y II (n = 65)	W	Valor p
Validez	4.32 (n = 64)	3.32 (n = 65)	3090	<0.01
Fiabilidad	4.55 (n = 62)	3.29 (n = 65)	3166	<0.01
Eficiencia	4.49 (n = 61)	2.48 (n = 65)	3533.5	<0.01
Retroalimentación	3.98 (n = 65)	2.56 (n = 64)	3298.5	<0.01
Interpretación	3.78 (n = 65)	2.86 (n = 65)	2983.5	<0.01
Claridad	3.89 (n = 65)	2.86 (n = 65)	3079.5	<0.01
Metodología	4.41 (n = 58)	2.80 (n = 65)	3088.5	<0.01

COMPARATIVA ENTRE ASIGNATURAS DE TERCER AÑO

En el caso de tercer año, se comparó el ramo de Medicina Interna I y II asociado a la UNEV con dos que no lo estaban. En el primer caso al hacer la comparación con Fisiopatología I y II, sólo se observan diferencias significativas en las áreas de retroalimentación y metodología (p-value<0.01), esto a pesar de que los valores promedio de Medicina Interna superan a Fisiopatología en cinco de las siete dimensiones.

	Medicina Interna I y II (n = 59)	Fisiopatología I y II (n = 59)	W	Valor p
Validez	3.74 (n = 59)	3.77 (n = 57)	1721.5	0.82
Fiabilidad	4.30 (n = 53)	3.76 (n = 59)	1962	0.013
Eficiencia	3.54 (n = 59)	3.44 (n = 59)	1838.5	0.58
Retroalimentación	3.71 (n = 59)	2.61 (n = 59)	2515	<0.01
Interpretación	3.27 (n = 59)	3.27 (n = 59)	1730.5	0.95
Claridad	3.62 (n = 56)	3.13 (n = 59)	2007.5	0.04
Metodología	4.37 (n = 51)	3.72 (n = 59)	1966	<0.01

Al hacer la comparación con Farmacología I y II, se observan diferencias significativas en 5 de las dimensiones, pero en este caso presenta mejores indicadores el ramo de Farmacología, exceptuando en fiabilidad y metodología, donde si bien los promedios fueron mayores, no se obtuvo una mayor significancia al realizar pruebas de Mann-Whitney.

	Medicina Interna I y II (n = 59)	Farmacología I y II (n = 59)	W	Valor p
Validez	3.74 (n = 59)	4.42 (n = 55)	1122.5	<0.01
Fiabilidad	4.30 (n = 53)	4.52 (n = 52)	1170.5	0.14
Eficiencia	3.54 (n = 59)	4.61 (n = 56)	885	<0.01
Retroalimentación	3.71 (n = 59)	4.40 (n = 52)	1039.5	<0.01
Interpretación	3.27 (n = 59)	3.97 (n = 59)	1162	<0.01
Claridad	3.62 (n = 56)	4.34 (n = 50)	864	<0.01
Metodología	4.37 (n = 51)	4.58 (n = 55)	1176	0.10



Finalmente, en el caso de 5to año, no se observan diferencias significativas en ninguna de las áreas. EMQ-II presenta 4 valores promedio mayores, mientras que Ginecología y Obstetricia, 3, heterogeneidad que hace más compleja la comparación entre ambos ramos respecto al impacto de la UNEV.

	Ginecología y Obstetricia (n = 10)	EMQ-II (n = 16)	W	Valor p
Validez	3.4 (n = 10)	3.81 (n = 16)	71.5	0.63
Fiabilidad	4.37 (n = 8)	3.93 (n = 15)	80.5	0.13
Eficiencia	4.33 (n = 9)	3.93 (n = 16)	85.5	0.40
Retroalimentación	2.7 (n = 10)	3.25 (n = 16)	217	0.49
Interpretación	2.9 (n = 10)	4.27 (n = 11)	27	0.03
Claridad	3.1 (n = 10)	3.56 (n = 16)	66	0.44
Metodología	3.6 (n = 10)	3.31 (n = 16)	93	0.48

COMPARATIVA ENTRE ASIGNATURAS DE QUINTO AÑO

COMPARACIONES

SEGÚN VARIABLES DE SEGMENTACIÓN



SEMIOLOGÍA I Y II

- **Tercer año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Cuarto año en relación al total**

Se encontraron diferencias en claridad, con un valor-p de 0.0079.

- **Quinto año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Tercer año en relación a cuarto año**

Se encontraron diferencias en cuanto a retroalimentación y claridad, con un valor-p de 0.0057 y 0.0012 respectivamente.

- **Tercer año en relación a quinto año**

No se encontraron diferencias significativas entre estos niveles en ninguna de las siete categorías analizadas en el estudio.

- **Cuarto año en relación a quinto año**

Se encontraron diferencias en cuanto a validez, claridad y metodología, con un valor-p de 0.0037, 0.0048 y 0.00062, respectivamente.

FISIOLOGÍA I Y II

- **Tercer año en relación al total**

Se encontraron diferencias en cuanto a metodología, con un valor-p de 0.00012.

- **Cuarto año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Quinto año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Tercer año en relación a cuarto año**

Se encontraron diferencias en cuanto a validez y metodología, con un valor-p de 0.0011 y 0.00001, respectivamente.

- **Tercer año en relación a quinto año**

Se encontraron diferencias en cuanto a validez, eficiencia y metodología, con un valor-p de 0.00045, 0.00018 y 0.00001, respectivamente.

- **Cuarto año en relación a quinto año**

No se encontraron diferencias significativas entre estos niveles en ninguna de las siete categorías analizadas en el estudio.

MEDICINA INTERNA I Y II

- **Tercer año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Cuarto año en relación al total**

Se encontraron diferencias en cuanto a validez y retroalimentación, con un valor-p de 0.003391 y 0.0024, respectivamente.

- **Quinto año en relación al total**

Se encontraron diferencias en metodología, con un valor-p de 0.0002.

- **Tercer año en relación a cuarto año**

Se encontraron diferencias en cuanto a validez, fiabilidad y retroalimentación, con un valor-p de 0.009, 0.0008 y 0.002, respectivamente.

- **Tercer año en relación a quinto año**

No se encontraron diferencias significativas entre estos niveles en ninguna de las siete categorías analizadas en el estudio.

- **Cuarto año en relación a quinto año**

Se encontraron diferencias en validez ($p = 0.000019$), fiabilidad ($p = 0.0031$), retroalimentación ($p = 0.0000098$), interpretación ($p = 0.0089$), claridad ($p = 0.0026$) y metodología ($p = 0.000011$).

FISIOPATOLOGÍA I Y II

- **Tercer año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Cuarto año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Quinto año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Tercer año en relación a cuarto año**

No se encontraron diferencias significativas entre estos niveles en ninguna de las siete categorías analizadas en el estudio.

- **Tercer año en relación a quinto año**

Se encontraron diferencias en validez ($p = 0.0092$), eficiencia ($p = 0.0027$), retroalimentación ($p = 0.0051$), interpretación ($p = 0.0089$) y metodología ($p = 0.0028$).

- **Cuarto año en relación a quinto año**

No se encontraron diferencias significativas entre estos niveles en ninguna de las siete categorías analizadas en el estudio.

FARMACOLOGÍA I Y II

- **Tercer año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Cuarto año en relación al total**

No se encontraron diferencias significativas en ninguna de las siete categorías analizadas en el estudio.

- **Quinto año en relación al total**

Se encontraron diferencias en cuanto a metodología, con un valor-p de 0.0081.

- **Tercer año en relación a cuarto año**

No se encontraron diferencias significativas entre estos niveles en ninguna de las siete categorías analizadas en el estudio.

- **Tercer año en relación a quinto año**

Se encontraron diferencias en cuanto a metodología, con un valor-p de 0.00069.

- **Cuarto año en relación a quinto año**

Se encontraron diferencias en cuanto a metodología, con un valor-p de 0.0087.

RESULTADOS

Dos profesores encargados de curso (PEC) compartieron su experiencia con respecto a la Unidad de Evaluación (UNEV). Ambos piensan que ha sido de gran utilidad contar con una Unidad de Evaluación en la Escuela de Medicina. Una profesora destaca:

> ***"Es una fortaleza de nuestra institución contar con una unidad que conoce sobre metodologías de evaluación, además de supervisar, gestionar y desarrollar estrategias para la evaluación con altos estándares de calidad y confiabilidad."***

Por otra parte, ambos docentes consideran que la calidad de sus evaluaciones ha mejorado desde que trabajan con la UNEV. La misma profesora señala:

> ***"Como docente, mi percepción es que la calidad de las evaluaciones ha mejorado en su redacción, en claridad y en imparcialidad."***

El otro profesor, que ha trabajado con la UNEV desde su creación, comienza explicando exactamente qué papel cumple esta:

> ***"Tienen instrucciones sobre cómo hacer las pruebas y se realizan sesiones de validación donde se revisan todas las preguntas. Luego nos informan las preguntas objetadas y nosotros consideramos si corresponde o no el cambio."***

Luego continúa, señalando sus motivos para consultar con ella. Uno de ellos es que gran parte de los docentes se benefician de aprender acerca de docencia con la ayuda de la UNEV. Por otra parte, la UNEV le ayuda a manejar los tiempos, dado que insiste en hacer las cosas a tiempo y bien. Finalmente, el docente señala que la UNEV es útil para ecualizar las pruebas entre sí:

> ***"Antes cada campus organizaba ciertas pruebas, y esto provocaba disparidad."***

DE ENTREVISTA A **DOCENTES**

BENCHMARKING

CASOS RELEVANTES EN OTRAS INSTITUCIONES EDUCATIVAS

Luego de analizar la UNEV de la Escuela de Medicina de la Universidad de Chile, investigamos acerca de otros entes que generan recomendaciones para la realización de pruebas en Escuelas de Medicina.

UNIVERSIDAD SAN SEBASTIAN

DEPARTAMENTO DE EDUCACIÓN MÉDICA

El Departamento de Educación Médica ofrece diversos recursos para docentes, incluyendo consejos para realizar clases, recomendaciones para evaluar los aprendizajes y orientaciones para el desarrollo de evaluaciones, entre otras. Con respecto a esto último, se otorgan orientaciones para el desarrollo de evaluaciones tanto en línea como presenciales. Además de contar con orientaciones oficiales para la confección de solemnes de selección múltiple, existe un protocolo de validación de solemnes, donde se realizan reuniones con la unidad y se envía el solemne validado para ser rendido por los estudiantes. Finalmente, se analizan los resultados para realizar un informe y reflexionar como cuerpo docente.

UNIVERSIDAD DE LOS ANDES

OFICINA DE EDUCACIÓN MÉDICA

La Oficina de Educación Médica es un equipo interdisciplinario que se dedica a la mejora continua de la educación médica en la Universidad. En su página ofrece diversos recursos, incluyendo normativas para confeccionar evaluaciones, pautas de evaluación, reglamentos de propiedad intelectual, protocolos ante situaciones inesperadas y tutoriales para el uso de tecnologías en clases, entre otros. Cabe destacar que la Oficina de Educación Médica no participa propiamente tal en la creación de evaluaciones como la UNEV, sino que otorga lineamientos y recomendaciones para ella.

UNIVERSIDAD DEL DESARROLLO

CENTRO DE DESARROLLO EDUCACIONAL DE LA FACULTAD DE MEDICINA

El objetivo del Centro de Desarrollo Educacional de la Facultad de Medicina es, según se indica en su página:

"Generar y comunicar orientaciones y estándares en relación a los procesos de enseñanza-aprendizaje y evaluación que ocurren en la Facultad"

Produce material educativo sobre Docencia Médica, evalúa los tutores clínicos de la facultad y asesora en la producción de instrumentos de evaluación, entre otros.

UNIVERSIDAD FINIS TERRAE

CENTRO DE INVESTIGACIÓN EN EDUCACIÓN MÉDICA Y CIENCIAS DE LA SALUD

En su página indica que su misión es:

"Abrirse a la comunidad de docentes y estudiantes de carreras de ciencias de la salud para compartir recursos pedagógicos que colaboren en el fortalecimiento de la educación en ciencias de la salud"

Dentro de sus trabajos se encuentran las "orientaciones para el desarrollo de evaluaciones por vía remota para académicos". Además de poseer varios recursos docentes, incluyendo una serie de videos y boletines con el fin de fortalecer la educación en las ciencias de la salud. Uno de los boletines que destacamos es las cápsulas de evaluación de los aprendizajes.



DISCUSIÓN

En la carrera de Medicina en la Universidad de Chile el proceso de estandarización es llevado a cabo por la Unidad de Evaluación (UNEV). Este boletín se centró en analizar el impacto que esta tiene sobre la calidad de las evaluaciones de selección múltiple de esta carrera, considerando tanto a alumnos como a docentes.

ENCUESTA A ALUMNOS

Tras el análisis de los resultados, se puede concluir que en segundo año los promedios de Semiología son mayores a los de Fisiología en todas las áreas con resultados significativos. Esto puede ser atribuido a la UNEV, pero también puede deberse a otras diferencias fundamentales entre ambos ramos, como la naturaleza clínica de Semiología y los cambios de programa de Fisiología.

Dentro de tercer nivel, Medicina Interna sólamente es significativamente superior a Fisiopatología en las áreas de "retroalimentación" y "metodología". En cuanto a la comparación con Farmacología, este supera significativamente a Medicina Interna en casi todos los puntos.

En cuanto a quinto nivel, las diferencias entre ramos no son significativas y consideramos que la cantidad de respuestas que se obtuvieron no son suficientes para llevar a cabo un análisis concluyente.

Por otra parte, se realizó análisis según el año que los alumnos están cursando actualmente. Se evidenció respuestas diferentes dependiendo de en qué año se habían rendido los ramos, lo cual se le puede atribuir principalmente a cambios en cuanto a programa (principalmente en Fisiología), cambios en cuanto al equipo docente (evidenciado principalmente en Fisiopatología) y cambios en modalidad (ramos clínicos como Medicina Interna y Semiología).

Consideramos que el impacto de la UNEV no es claramente percibido por los estudiantes, dado que no existe una diferencia clara entre aquellas asignaturas que reciben asesoría de esta y aquellas que no. Por otra parte, creemos que existen diferencias fundamentales entre las diferentes asignaturas que no permiten llevar a cabo un análisis claro del impacto de la UNEV; para esto habría que conocer la percepción de los alumnos previo a la UNEV y posterior a ella.

ENTREVISTA A DOCENTES

Luego de entrevistar a los profesores, se evidenció que la UNEV ha tenido un impacto positivo sobre su trabajo. Los profesores consideran que esta les ha permitido mejorar la calidad de sus evaluaciones, aprender cosas nuevas en el campo de la educación médica, manejar mejor los tiempos para el desarrollo de evaluaciones y ecualizar las pruebas entre sí. Los académicos destacan la importancia de la UNEV y consideran que maneja altos estándares de calidad.

Si bien no es posible afirmar con certeza que los alumnos perciben el impacto de la UNEV, es claro que los profesores encargados de curso entrevistados sí lo perciben.

LIMITACIONES

Dado que la encuesta aplicada al alumnado es de carácter voluntario, la principal limitación es la participación del alumnado para responder la encuesta, y por tanto, reunir una cantidad suficiente de respuestas para lograr una muestra confiable de opiniones para el análisis posterior se tornó un desafío difícil de abordar. Esto es particularmente relevante en el caso de sexto y séptimo año, donde, pese a haber difusión dentro de sus grupos de generación, la participación fue nula. La participación estudiantil en la carrera de Medicina ha ido en caída, lo que se evidencia en la poca participación en las votaciones de los centro de alumnos de la carrera, la poca participación en las evaluaciones docentes, en actividades voluntarias de la carrera y de la facultad, donde se ve una menor cantidad de alumnos de la cohorte siendo parte de estas.

Por otra parte, no existen ramos de cuarto año que sean asesorados por la UNEV, impidiendo realizar un análisis de las asignaturas de este nivel.

Finalmente, una limitación importante es que sólo fue posible entrevistar a 2 profesores encargados de curso para el boletín, dado que fue difícil ponerse en contacto con más docentes y solicitar su participación.

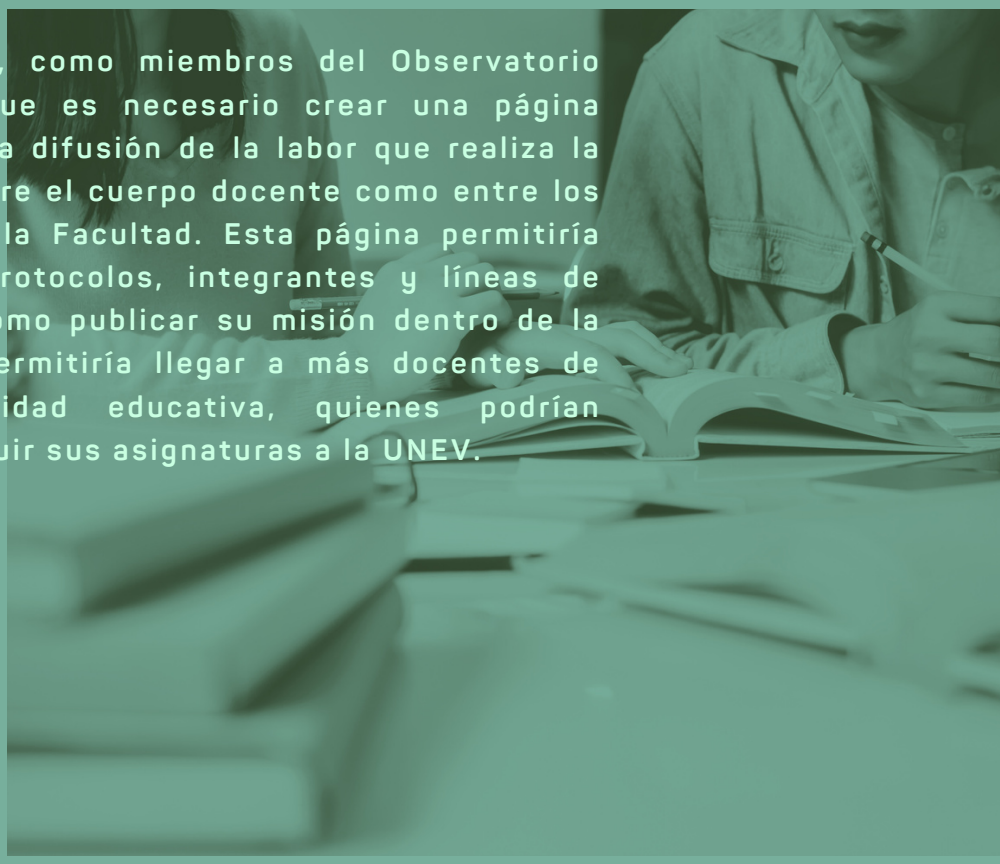


RECOMENDACIONES

Los resultados de la encuesta al alumnado no fueron concluyentes, pero las apreciaciones de los docentes sobre la labor realizada por la UNEV fueron positivas, destacando su metodología y sus estándares de calidad. Consideramos que los profesores encargados de curso son quienes perciben en mayor medida el trabajo realizado por la UNEV, dado que estuvieron presentes antes y después de dicho trabajo. Creemos que la mayoría de los ramos se beneficiarían de la asesoría entregada por la UNEV, permitiendo elevar la calidad de las evaluaciones en la carrera de Medicina en la Universidad de Chile.

Por otra parte, durante el proceso de Benchmarking quedó en evidencia la falta de recursos disponibles en línea que den a conocer la labor de la UNEV, muestren los protocolos utilizados u otorguen una línea de contacto para trabajar con ella. En contraste, varias universidades cuentan con una página para estos propósitos, como el Centro de Desarrollo Educacional de la Universidad del Desarrollo.

En ese sentido, como miembros del Observatorio consideramos que es necesario crear una página para aumentar la difusión de la labor que realiza la UNEV, tanto entre el cuerpo docente como entre los estudiantes de la Facultad. Esta página permitiría transparentar protocolos, integrantes y líneas de contacto, así como publicar su misión dentro de la carrera. Esto permitiría llegar a más docentes de nuestra comunidad educativa, quienes podrían motivarse a incluir sus asignaturas a la UNEV.



CONCLUSIÓN

Luego del trabajo investigativo realizado, queremos destacar el rol de la Unidad de Evaluación y esperamos que aumente el conocimiento de ella entre los académicos y alumnos de la carrera de Medicina.

Si bien la UNEV es relativamente desconocida y no existen muchos profesores que hayan trabajado con ella, creemos que su labor es destacable y que tiene el potencial de mejorar la calidad de las evaluaciones y facilitar la tarea de los profesores en la carrera de Medicina.



BIBLIOGRAFÍA

- Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2015). How effective are selection methods in medical education? A systematic review. Medical Education, 50(1), 36-60. doi:10.1111/medu.12817
- Epstein, R. M. (2007). Assessment in Medical Education. New England Journal of Medicine, 356(4), 387-396. doi:10.1056/nejmra054784
- SEPAUC. (2017). ¿Qué debemos considerar al aplicar una prueba estandarizada? Sociedad de Psicólogos y Psiquiatras de la Universidad de Chile. Recuperado el 9 de mayo de 2023, de https://sepauc.cl/que-debemos-considerar-al-aplicar-una-prueba-estandarizada/#:~:text=Las%20evaluaciones%20estandarizadas%20se%20denominan,hace%20de%20los%20datos%20obtenidos.
- Hernández, M., Ramírez, E., & Gamboa, S. (2018). La implementación de una evaluación estandarizada en una institución de educación superior. Revista de la Educación Superior, 47(185), 149-169. Recuperado de https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-26732018000100149
- Sandoval Rubilar, Pedro, Maldonado-Fuentes, A. C., Tapia-Ladino, M., Sandoval Rubilar, Pedro, Maldonado-Fuentes, A. C., & Tapia-Ladino, M. (2022). Evaluación educativa de los aprendizajes: Conceptualizaciones básicas de un lenguaje profesional para su comprensión. Páginas de Educación, 15(1), 49–75. https://doi.org/10.22235/pe.v15i1.2638
- Villa, A., Galindo-Domínguez, H., Loaiza, M. I., & Andrade, P. (2022). Percepción valorativa de los estudiantes sobre la calidad del aprendizaje en sus estudios. Profesorado, Revista De Currículum Y Formación Del Profesorado, 26(1), 41–63. https://doi.org/10.30827/profesorado.v26i1.13727

ANEXOS

- Consentimiento Informado para Participación en la Encuesta:

Estimados participantes,

Están siendo invitados a participar en una investigación sobre Percepción sobre Metodologías de Evaluación de asignaturas impartidas en la carrera de Medicina de la Universidad de Chile. Su participación es voluntaria y su colaboración es importante para el éxito de este estudio.

Propósito del Estudio

El propósito de este estudio es comparar la percepción de estudiantes de medicina de la Universidad de Chile de segundo a séptimo año con respecto a asignaturas cuyas evaluaciones se rigen según la Unidad de Evaluación de la Escuela de Medicina (UNEV), versus aquellas que no siguen sus protocolos. Sus respuestas ayudarán a crear recomendaciones para los distintos Profesores Encargados de Curso, así como también para la Escuela de Medicina con respecto a la eficacia de la UNEV.

Procedimiento

Si decide participar, se le pedirá que complete una encuesta en línea a través de Google Forms. La encuesta incluirá preguntas sobre su experiencia en relación a las evaluaciones de opción múltiple de la asignatura. Se estima que tomará aproximadamente 7-10 minutos completar la encuesta.

Confidencialidad

Todas las respuestas serán tratadas de manera confidencial. Sus respuestas serán anónimas y no se compartirán datos individuales en ningún informe o publicación resultante. La información recopilada se utilizará únicamente para fines de investigación y no se utilizará para ningún otro propósito.

Voluntariedad

Su participación en este estudio es completamente voluntaria. Puede elegir no participar o retirarse en cualquier momento sin ninguna consecuencia. No habrá ninguna penalización si decide no participar o retirarse.

Contacto

Si tiene alguna pregunta sobre este estudio antes de decidir participar, puede comunicarse directamente con Nicolás A. Lavados, nicolaslavados@ug.uchile.cl.

Consentimiento

Al hacer clic en el enlace y completar la encuesta, está indicando que ha leído la información proporcionada, que entiende los detalles del estudio y que da su consentimiento para participar.

Por favor, tómese un momento para leer y comprender la información proporcionada anteriormente antes de decidir si desea participar en este estudio.

Al hacer clic en el enlace y completar la encuesta, está indicando que ha leído la información proporcionada, que entiende los detalles del estudio y que da su consentimiento para participar.

- Consentimiento Informado para Participación de Docentes:

Consentimiento Informado

Estimado(a) docente:

Está siendo invitado(a) a participar en una investigación sobre el impacto de la Unidad de Evaluación de la Escuela de Medicina (UNEV). Su participación es voluntaria y su colaboración es importante para el éxito de este estudio.

El propósito de este estudio es conocer la percepción de los estudiantes y los profesores sobre la UNEV, tal que sea posible generar recomendaciones dirigidas a los distintos Profesores Encargados de Curso y a la Escuela con respecto a la incorporación de otras asignaturas a la UNEV.

Todas las respuestas serán tratadas de manera confidencial. Estas serán anónimas y no se compartirán datos individuales en ningún informe o publicación resultante. La información recopilada se utilizará únicamente para fines de investigación y no se usará para ningún otro propósito.

Su participación en este estudio es completamente voluntaria. Puede elegir no participar o retirarse en cualquier momento sin ninguna consecuencia. No habrá ninguna penalización si decide no participar o retirarse.

Al firmar este consentimiento informado, está indicando que ha leído la información proporcionada, que entiende los detalles del estudio y que da su consentimiento para participar.

Por favor, tómese un momento para leer y comprender la información proporcionada anteriormente antes de decidir si desea participar en este estudio.
Al hacer clic en el enlace y completar la encuesta, está indicando que ha leído la información proporcionada, que entiende los detalles del estudio y que da su consentimiento para participar.

Contacto:

Si tiene alguna pregunta sobre este estudio antes de decidir participar, puede comunicarse directamente con Nicolás A. Lavados, nicolaslavados@ug.uchile.cl. Muchas gracias por su participación.

_____
Firma de docente

- Cálculo de población de estudio

La población total de estudio en la escuela de medicina es de 1708 estudiantes, de los cuales se elimina a primer y segundo año, debido a que a la fecha de la investigación no se encuentran sin haber completado el ramo clínico de semiología II. Además se eliminan sexto y séptimo año por falta de participación y por la cantidad de ramos que se dieron en tiempos de pandemia que pueden afectar su percepción de los ramos. Por lo cual nos queda como población los estudiantes de tercero, cuarto y quinto, los cuales según la información a partir de las personas inscritas en las asignaturas medicina interna II, cirugía, geriatria, pediatria y ginecologia, esto se debe a que estos ramos son los que poseen prerrequisitos necesarios para poder validar de manera acertiva la presencia de un estudiante en tercer, cuarto y quinto año. Por lo tanto, la población fue de 709 estudiantes.

CURSO	Nº DE PARTICIPANTES
ME06038 Medicina Interna II	81
ME06038 Medicina Interna II	40
ME06038 Medicina Interna II	41
ME06038 Medicina Interna II	40
ME06038 Medicina Interna II	34
ME07041 Cirugía	21
ME07041 Cirugía	18
ME07041 Cirugía	20
ME07041 Cirugía	20
ME07041 Cirugía	20
ME07043 Geriatría	121
ME09056 Pediatría y Cirugía Infantil	28
ME09056 Pediatría y Cirugía Infantil	41
ME09056 Pediatría y Cirugía Infantil	20
ME09056 Pediatría y Cirugía Infantil	22
ME09056 Pediatría y Cirugía Infantil	16
ME10060 Ginecología-Obstetricia	28
ME10060 Ginecología-Obstetricia	26
ME10060 Ginecología-Obstetricia	28
ME10060 Ginecología-Obstetricia	24
ME10060 Ginecología-Obstetricia	20

Para el cálculo del tamaño de muestra, se utilizó un nivel de confianza del 95%, y un margen de error del 5%, con una desviación estándar de 0,5 sacada de un estudio descriptivo de similares características. Así, se determinó que, para obtener resultados confiables, se requiere una muestra mínima de 249 estudiantes. Esto se calculó utilizando la fórmula del tamaño de muestra para muestreo estratificado en poblaciones finitas.



OBSERVATORIO DE
EDUCACIÓN MÉDICA

UNIVERSIDAD DE CHILE