

Validez y confiabilidad del instrumento de evaluación docente en las residencias médicas de la Universidad Nacional Autónoma de México

Validity and reliability of the teaching evaluation instrument in medical residencies at National Autonomous University of Mexico

Liz Hamui-Sutton-Sutton¹ , Ruth Fuentes-García² , Ivonne Zepeda-Gutiérrez¹ 

Resumen

Introducción: La evaluación docente (ED) permite a los residentes apreciar la calidad de la enseñanza recibida en su formación profesional. Los cuestionarios de opinión son una herramienta efectiva para identificar fortalezas y áreas de mejora de los docentes. La validez y la confiabilidad de estos instrumentos son fundamentales para asegurar la precisión y utilidad de los datos. **Materiales y métodos:** Se rediseñó el instrumento de ED utilizado en el Plan Único de Especializaciones Médicas, consta de 6 apartados, 4 dimensiones y se añadió una sección sobre el tiempo que dedican los profesores a la clínica, a las clases en el aula y a las tutorías. En el ciclo 2022-2023 se generaron 32.294 cuestionarios que fueron considerados en los análisis estadísticos realizados. **Resultados:** El tiempo promedio dedicado por los profesores a actividades clínicas fue del 48,35%, en el aula 25,33% y en las tutorías 26,32%. El coeficiente alfa de Cronbach fue 0,99 y el análisis factorial confirmatorio identificó cuatro factores que explican el 89,3% de la variabilidad del cuestionario. **Discusión:** Los resultados indicaron que los profesores fueron mejor evaluados en profesionalismo, actualización de conocimientos y calidad de la evaluación del aprendizaje, en cambio, el uso de tecnología educativa fue identificado como área de oportunidad. Además, se encontraron diferencias significativas según género, las docentes mujeres resultaron mejor evaluadas que los hombres; por el tipo de nombramiento, pues los profesores titulares fueron mejor calificados que los adjuntos; y por especialidades con evaluaciones más altas en las que combinan lo médico con lo quirúrgico.

Palabras clave: evaluación docente; validez; confiabilidad; residencias médicas; competencias docentes; práctica clínica

Abstract

Introduction: Teaching evaluation (TE) allows residents to appreciate the quality of the teaching received in their professional training. Opinion questionnaires are an effective tool for identifying strengths and areas of improvement for teachers. The validity and reliability of these instruments are relevant to ensuring the accuracy and usefulness of the data. **Materials and methods:** The TE instrument used in the *Plan Único de Especializaciones Médicas* was redesigned, consisting of 6 parts, 4 dimensions, and a section about the time that teachers dedicate to coaching at the clinical practice, teaching in the classroom, and mentoring residents. In the 2022–2023 cycle, 32,294 questionnaires were generated and were considered in the statistical analyses carried out. **Results:** The average time dedicated by teachers to clinical activities was 48.35%, in class 25.33%, and mentoring 26.32%. Cronbach's alpha coefficient was 0.99, and the confirmatory factor analysis identified four factors that explained 89.3% of the questionnaire's variability. **Discussion:** The results indicate that teachers were better evaluated in terms of professionalism, knowledge updating, and learning evaluation. In addition, educational technology was identified as an area of opportunity. Furthermore, significant differences were found according to gender; female teachers were better evaluated than men; due to the type of appointment, tenured professors were better qualified than assistant professors; and specialties with the highest evaluations were those that combine medical and surgical characteristics.

Keywords: teaching evaluation; validity; reliability; medical residencies; teaching competencies; clinical practice.

Fecha de envío: 2024-10-21 - Fecha de aprobación: 2024-12-02

(1) División de Estudios de Posgrado. Facultad de Medicina. Universidad Nacional Autónoma de México. Ciudad de México. México.

(2) Departamento de Matemáticas. Facultad de Ciencias. Universidad Nacional Autónoma de México. Ciudad de México. México.

*Autor de correspondencia: lizhamui@gmail.com



Introducción

En las residencias médicas, la evaluación docente (ED) es relevante, pues los docentes son un elemento clave en la formación de los futuros especialistas. Estos profesores guían y supervisan a los residentes, ayudándoles a adquirir las competencias necesarias en la atención al paciente y moldeando su desarrollo profesional y ético (Hamui *et al.*, 2024). Por ello, la retroalimentación que reciben en la ED es crucial para su crecimiento académico y profesional. Existen diversas formas de realizar la ED, como el portafolio docente, la autoevaluación, la evaluación por pares y la evaluación por parte de los estudiantes (Celis *et al.*, 2018). Los cuestionarios de opinión aplicados a los médicos residentes al final de cada ciclo escolar son una de las formas más accesibles y sencillas para que los docentes identifiquen sus fortalezas y debilidades (Hamui-Sutton, 2022), contribuyendo así, a la mejora continua de los programas de especialidades.

En este contexto, los cuestionarios de ED deben tener características métricas robustas como la validez y la confiabilidad. La validez se refiere a qué tan bien una prueba mide lo que se propone medir, y la confiabilidad a la consistencia de los resultados obtenidos en diferentes momentos (Medina-Díaz & Verdejo-Carrión, 2020).

Verificar que los cuestionarios sean válidos y confiables asegura que la retroalimentación a los docentes sea precisa, justa y útil para su desarrollo profesional y que mejore la calidad de la enseñanza impartida a los residentes (Luchsinger *et al.*, 2023). Para su efectividad, se recomienda que su aplicación sea continua, periódica, realizada fuera del estrés de las tareas cotidianas y de manera anónima.

La División de Estudios de Posgrado (DEP) de la Facultad de Medicina (FM) de la UNAM ha llevado a cabo evaluaciones anuales desde 2014 para mejorar la calidad del desempeño educativo de los profesores de especialidades médicas. El Plan Único de Especializaciones Médicas (PUEM) describe el modelo pedagógico establecido por consenso entre organizaciones de salud y universidades para definir el currículo de cada especialidad (PUEM, DEP, FM, UNAM). Hasta el ciclo 2021-2022, se usó un instrumento con tres dimensiones: interacciones académicas, supervisión académica y profesionalismo, aplicado en línea. El cuestionario fue analizado para medir su validez y confiabilidad, las cifras obtenidas fueron 0,984 en el Alpha de Cronbach y una varianza explicada de 74,47%, 81,04 y 85,97 en cada una de las dimensiones respectivamente (Hamui-Sutton *et al.*, 2017). En la Figura 1 se muestran los promedios de calificación obtenidos por los profesores en los ocho ciclos de 2014-2015 a 2021-2022.

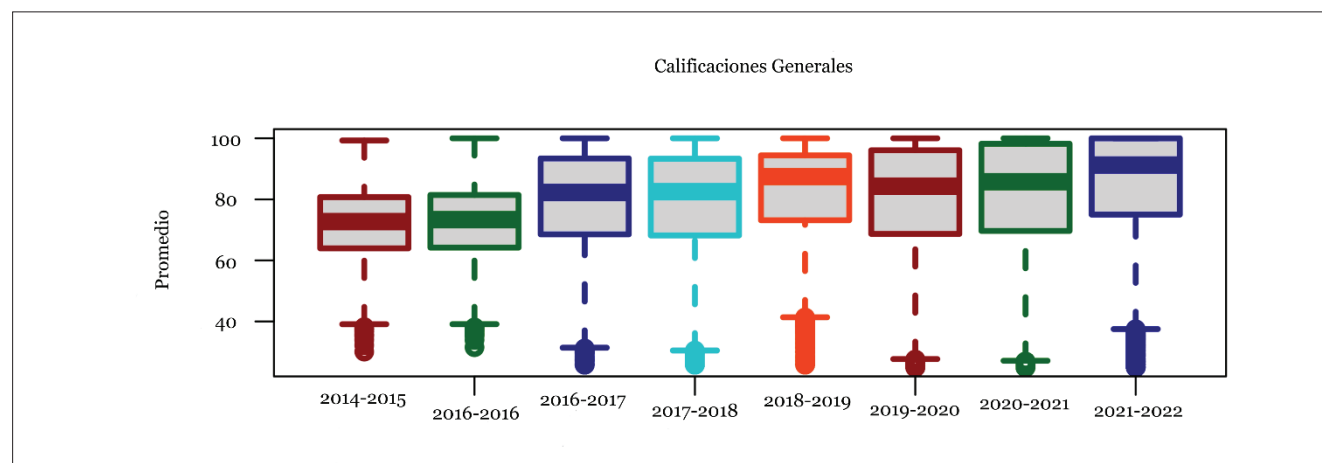


Figura 1: Diagrama de caja con el promedio por año de la ED basado en la primera versión del cuestionario. Se muestran los promedios obtenidos por los profesores en los ocho ciclos de 2014-15 a 2021-22, evaluando desde 0 a 100 por ciento (Hamui-Sutton *et al.*, 2023).

En 2022 se modificó el cuestionario, ahora con una sección inicial donde los residentes describen en porcentajes el tiempo que sus profesores dedican a actividades clínicas, en el aula y tutorías, reorganizando las secciones en cuatro dimensiones. En el apartado de resultados se describe en detalle esta versión del instrumento. A partir de los datos obtenidos del cuestionario, surgen las siguientes preguntas de investigación:

1. ¿Qué tan válido y confiable es el nuevo instrumento de ED respondido por los residentes del PUEM al final del ciclo escolar 2022-2023?
2. ¿Qué aporta esta nueva versión del instrumento de ED para las residencias médicas?
3. ¿Qué competencias docentes fueron las mejor evaluadas por los residentes y en qué aspectos hay áreas de oportunidad?

El cuestionario de ED se basa en competencias docentes, no en atributos personales del profesor. La educación basada en competencias (EBC) sigue el modelo de CanMEDS, desarrollado por el Royal College of Physicians and Surgeons of Canada, que incluye siete roles principales del médico: experto médico, comunicador, colaborador, líder, promotor de la salud, académico y profesional (Frank *et al.*, 2015). En la ED, se destaca el rol académico, aunque en la práctica real, los profesores muestran simultáneamente todos los roles mencionados.

Material y métodos

En 2022 se rediseñó el instrumento de ED con el fin de incluir no sólo actividades docentes en la clínica, sino también las que realizan en el aula y la tutoría. El objetivo de este artículo es describir estadísticamente la validez y confiabilidad estadística del cuestionario de ED, y analizar las diferencias y aportaciones de la nueva versión.

El cuestionario consta de 6 secciones: 1) actividades académicas (15 preguntas), 2) profesionalismo (9 preguntas), 3) docencia en la clínica (4 preguntas), 4) docencia en el aula (10 preguntas), 5) preguntas de opinión sobre la satisfacción con la sede y el cumplimiento del PUEM (3 preguntas) y 6) pregunta abierta para comentarios escritos sobre el desempeño del docente. En total, el instrumento consta de 41 reactivos con opciones de respuesta en escalas tipo Likert. En las secciones 3 y 4, se agregó la opción “no aplica” considerando esta posibilidad en las especialidades no clínicas; en los demás apartados esta opción fue eliminada. Las residencias que no incluyen actividad clínica como patología o medicina forense encuentran en esta versión del cuestionario la opción de no aplica para evitar introducir sesgos en las respuestas pues no tienen práctica clínica.

Las calificaciones promedio para los profesores se obtienen considerando en cada sección las respuestas válidas plasmadas en las escalas tipo Likert se utilizaron para construir un puntaje aditivo en donde todos los ítems tienen el mismo valor. Se obtuvo un promedio de las evaluaciones de todos los residentes inscritos en el curso que participaron en la evaluación, escalado a cantidades porcentuales.

Para contextualizar con algunos datos sobre la población a la que se aplicó el instrumento, en el ciclo académico 2022-2023 se registraron 14.695 médicas y médicos residentes en 79 especializaciones médicas. Se contabilizaron 712 cursos en 162 sedes hospitalarias avaladas por la FM de la UNAM. La edad promedio de los residentes fue de treinta años. De los 14.695 residentes, 8.195 (55,76%) fueron mujeres (Numeralia-DEP, 2024).

El cuestionario fue respaldado por la DEP de la FM de la UNAM y contó con el apoyo del Departamento de Cómputo y Telecomunicaciones para la operación del sistema de ED. En el ciclo 2022-2023, 9.779 de 11.199 (87,32%) residentes evaluaron a 767 de 853 (89,91%) profesores. Se abarcaron 76 especialidades de 79 (94,20%) y 82 de 96 (91,11%) sedes. En total se contabilizaron 32.294 cuestionarios, pues un residente podía completar más de un registro, ya que cuenta con más de un profesor (titular y adjuntos).

El cuestionario se aplica en línea, mediante una plataforma digital creada específicamente para ese fin. El proceso de implementación comienza con el registro al examen de conocimientos del PUEM, que presentan anualmente los residentes de todas las especialidades entre noviembre y enero. En la página web del registro se agrega el enlace URL de la ED. Al abrir el cuestionario, aparece el nombre del residente, su especialidad, sede y año que cursa. Como se observa en la Figura 2a, el cuestionario comienza con un apartado donde los residentes anotan en porcentajes el tiempo que cada uno de sus profesores dedica a las actividades en la clínica, en el aula y en la tutoría. Una vez desplegadas las preguntas, debajo de cada una aparece el nombre de hasta tres profesores titulares o adjuntos con nombramiento en la UNAM, así como las opciones de respuesta para su calificación (Figura 2b). Los datos se capturan en tiempo real en la base de datos de la DEP. Una vez depurada la base de datos, se utiliza para elaborar los reportes que consulta el profesorado de manera personalizada en el sistema. Estos datos también se analizan estadísticamente para identificar mejoras educativas, así como para realizar aportes en investigación.

a

Clasificación

Nombre del profesor	Profesor de Clínico	Profesor de Aula	Profesor de Tutor	Total	Observaciones
DRA. [REDACTED] MARÍA DE LA LUZ	0	0	0	<input type="text"/>	<input type="text"/>
DRA. [REDACTED] MERCEDES	0	0	0	<input type="text"/>	<input type="text"/>
DR. [REDACTED] JOSÉ NICOLÁS	0	0	0	<input type="text"/>	<input type="text"/>
DR. [REDACTED] MIGUEL ÁNGEL	0	0	0	<input type="text"/>	<input type="text"/>
DR. [REDACTED] HERNÁN	0	0	0	<input type="text"/>	<input type="text"/>

b

I. En la práctica el profesor interactúa con los residentes durante:

	Diario	Alguna vez en la semana	Alguna vez en el mes	Nunca
En la práctica el profesor interactúa con los residentes durante:	1. El paso de visita en hospitalización			
DRA. [REDACTED] MARÍA DE LA LUZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
DRA. [REDACTED] MERCEDES	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
En la práctica el profesor interactúa con los residentes durante:	2. Las sesiones de revisión de casos clínicos (expediente, estudios de gabinete y laboratorio, anatomopatológicos, morbimortalidad, etc.)			
DRA. [REDACTED] MARÍA DE LA LUZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
DRA. [REDACTED] MERCEDES	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
En la práctica el profesor interactúa con los residentes durante:	3. La consulta externa			
DRA. [REDACTED] MARÍA DE LA LUZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
DRA. [REDACTED] MERCEDES	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
En la práctica el profesor interactúa	4. La práctica en las áreas diagnósticas y quirúrgicas			

Figura 2: Representación de formato del cuestionario de ED que responden los médicos residentes: a) clasificación del tipo de actividad que realiza el profesor y porcentaje de tiempo dedicado; b) ejemplo de preguntas por docente (sistema de evaluación docente del PUEM, DEP, FM, UNAM).

Todos los análisis fueron realizados con el software de uso libre R (<https://www.r-project.org/>), en particular con las funciones básicas y con su librería "lavaan". Para la validación del instrumento de medición, se calculó el coeficiente alfa de Cronbach utilizando la matriz de correlación policórica, se realizó un análisis factorial confirmatorio utilizando el método de rotación varimax, estimando por mínimos cuadrados ponderados, y se utilizó el coeficiente omega para demostrar la confiabilidad. Además, se revisó pregunta por pregunta del instrumento para identificar en los promedios de calificaciones globales las competencias mejor evaluadas y las áreas de oportunidad del profesorado en general.

Resultados

El tiempo promedio que el profesorado dedicó a actividades clínicas fue del 48,35%, en el aula 25,33% y en la tutoría 26,32%.

El coeficiente Alfa de Cronbach tuvo un valor de 0,99 y, adicionalmente, se obtuvo una lambda_6 de Guttman igual a 1, lo que significa que el instrumento tiene un alto grado de confiabilidad. Se realizó un análisis factorial confirmatorio, el cual reveló la existencia de 4 factores con las siguientes etiquetas: WLS1, WLS2, WLS3 y WLS4, los cuales se pueden apreciar en la Tabla 1. El primer componente explica el 29,3% de la varianza, el segundo el 23,3%, el tercero el 21,5% y el cuarto el 15,2%. En conjunto, explican el 89,3% de la variabilidad total, lo cual es una cifra adecuada. Es destacable que las dos primeras secciones del instrumento están agrupadas en el primer factor, que corresponde a las actividades académicas y el profesionalismo. El resto de las secciones corresponde a un factor por sección, que era el objetivo. Esto indica que hay una congruencia entre el pensamiento de la persona que responde y la situación evocada en la pregunta.

Tabla 1: Cargas factoriales por reactivo. Se muestran el número de los reactivos y la carga de cada uno en los cuatro factores identificados (Fuente: elaboración propia).

	WLS1	WLS2	WLS3	WLS4		WLS1	WLS2	WLS3	WLS4
1	0,617	0,491	0,263	0,348	20	0,471	0,689	0,365	0,294
2	0,660	0,393	0,351	0,399	21	0,464	0,697	0,362	0,308
3	0,647	0,462	0,346	0,365	22	0,488	0,662	0,393	0,316
4	0,657	0,464	0,364	0,351	23	0,473	0,670	0,400	0,305
5	0,670	0,456	0,366	0,353	24	0,499	0,617	0,424	0,323
6	0,715	0,399	0,327	0,289	25	0,300	0,300	0,351	0,739
7	0,746	0,361	0,348	0,274	26	0,353	0,324	0,414	0,690
8	0,747	0,334	0,361	0,255	27	0,311	0,307	0,379	0,715
9	0,740	0,341	0,387	0,265	28	0,340	0,348	0,413	0,689
10	0,727	0,370	0,378	0,263	29	0,377	0,217	0,632	0,488
11	0,640	0,500	0,388	0,307	30	0,386	0,285	0,693	0,379
12	0,682	0,428	0,425	0,294	31	0,418	0,366	0,668	0,381
13	0,650	0,475	0,416	0,303	32	0,473	0,362	0,663	0,353
14	0,654	0,464	0,414	0,324	33	0,400	0,451	0,651	0,343
15	0,649	0,469	0,422	0,307	34	0,411	0,465	0,641	0,340
16	0,482	0,680	0,332	0,319	35	0,419	0,420	0,653	0,373
17	0,503	0,664	0,351	0,314	36	0,425	0,430	0,645	0,362
18	0,396	0,725	0,319	0,297	37	0,402	0,382	0,653	0,391
19	0,466	0,692	0,345	0,313	38	0,436	0,445	0,632	0,343

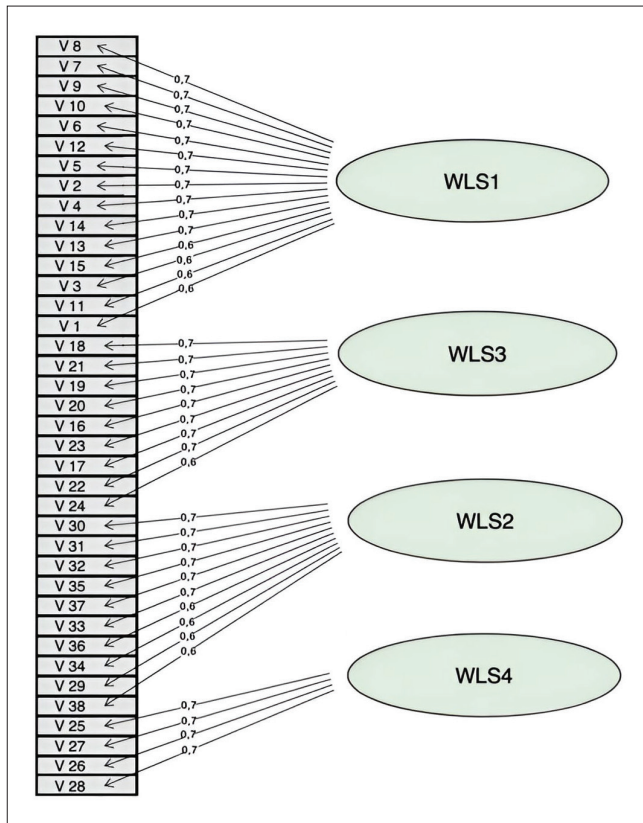


Figura 3: Agrupamiento de variables por factores en el cual los reactivos coinciden con las secciones del cuestionario. El primer factor agrupa las secciones uno y dos (Fuente: elaboración propia).

También se llevó a cabo el cálculo del coeficiente omega para realizar un análisis factorial confirmatorio con los cuatro factores mencionados anteriormente. A diferencia del coeficiente alfa, el coeficiente omega considera las cargas factoriales (Anderson & Gerbing, 1988). Los resultados obtenidos fueron los siguientes: primer factor: 0,94; segundo factor: 0,91; tercer factor: 0,91; cuarto factor: 0,86. Es decir, el factor con mayor confiabilidad fue el primero que agrupa las preguntas de las secciones 1) actividades académicas y 2) profesionalismo.

Se evaluaron los promedios generales por género de los profesores. La mediana del promedio para los hombres fue de 84,33 y para las mujeres de 86,02, y la diferencia en las medianas fue estadísticamente significativa (prueba de Wilcoxon, p-valor = 0,016) (Figura 4a). Se calcularon las medianas de la calificación promedio para los diferentes tipos de nombramiento. Se encontró que los profesores titulares tuvieron una mediana de 85,92, mientras que los adjuntos de 84,57. Se realizó una prueba de Wilcoxon para comparar estas medianas, obteniendo un p-valor de $1,235e-1$, lo que indica que la diferencia es estadísticamente significativa. Es decir, los profesores titulares obtuvieron una mejor evaluación en comparación con los adjuntos (Figura 4b).

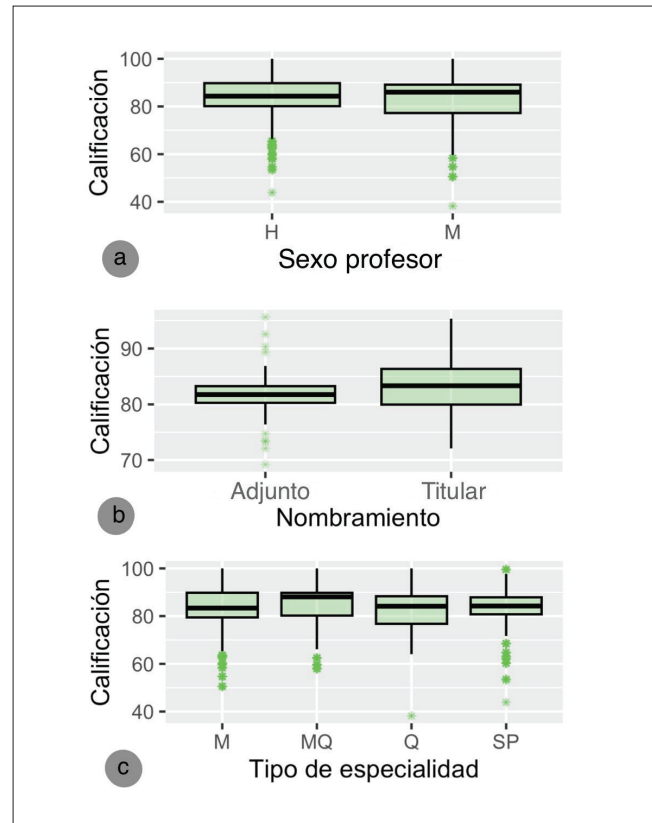


Figura 4: Promedio de calificaciones del profesorado del PUEM por: a) sexo, b) tipo de nombramiento y c) tipo de especialidad (M: médica; MQ: médico quirúrgica; Q: quirúrgica; SP: sin paciente) (Fuente: elaboración propia).

Se clasificaron las especialidades médicas en cuatro grupos (Anexo 1) y se calcularon las medianas de la calificación por grupo. Los resultados fueron los siguientes: médicas 83,38, médico-quirúrgicas 88,06, quirúrgicas 84,14 y sin pacientes 84,23. Se aplicó la prueba de suma de rangos de Kruskal-Wallis para comparar estas medianas y se obtuvo un p-valor menor a $2,2e-16$, lo que indica que las medianas son estadísticamente diferentes entre los grupos de especialidades médicas (Figura 4c).

Por otra parte, se examinaron los promedios de la calificación global por pregunta para valorar las competencias del profesorado. En la Tabla 2 se presentan los reactivos por sección del cuestionario para relacionarlos con las competencias de los siete roles del médico. Se indica el promedio global de cada pregunta considerando a los 767 profesores evaluados. De acuerdo con el modelo CanMEDS (Pritchard et al., 2023), el 50% de las preguntas (19) corresponden al rol de académico, lo que resulta lógico en un instrumento de ED. El 13% de los reactivos (5) se refieren al rol de experto y el 13% (5) al de comunicador. El 10,5% (4) se relacionan con el rol profesional, el 8% (3) con el de colaborador, el 5% (2) con el rol de líder y no hay preguntas que se relacionen con el rol de promotor de la salud.

Tabla 2: Promedio de calificación global del profesorado por reactivo y su relación con las competencias del modelo CanMeds.

Sección/ pregunta	1. Actividades académicas Opciones de respuesta: siempre (4), casi siempre (3), casi nunca (2), nunca (1)	Competencia modelo CanMeds. Médico como	Promedio calificación global del profesorado
1	Está actualizado en los conocimientos de su especialidad	Experto	0,87
2	Destina tiempo suficiente a la enseñanza y la supervisión	Académico	0,82
3	Plantea preguntas constructivas en el momento adecuado para motivar el aprendizaje	Académico	0,84
4	Identifica y aprovecha las oportunidades para enseñar o corregir con el fin de estimular el desarrollo académico del residente	Académico	0,83
5	Transmite la información necesaria para aplicarla en la práctica clínica, la docencia o la investigación	Comunicador	0,84
6	Promueve la investigación	Académico	0,83
7	Tiene destrezas tecnológicas para enseñar a distancia	Académico	0,81
8	Utiliza recursos virtuales para impartir su asignatura (videoconferencias, aulas virtuales, etc.)	Académico	0,80
9	Utiliza materiales digitales para preparar sus clases (como videos, infografías, plataformas educativas)	Académico	0,80
10	Interactúa a distancia con otras instancias educativas para fortalecer los aprendizajes	Académico	0,79
11	Es congruente entre lo que enseña y lo que evalúa	Académico	0,84
12	Emplea diferentes recursos o procedimientos para evaluar el curso	Académico	0,82
13	Realimenta y reflexiona acerca de las experiencias de aprendizaje de los residentes	Comunicador	0,82
14	Asesora y orienta el estudio	Comunicador	0,82
15	Coordina al equipo de salud del servicio para propiciar la enseñanza a los residentes	Colaborador	0,83
Sección/ pregunta	2. Profesionalismo Opciones de respuesta: Totalmente de acuerdo (4), de acuerdo (3), en desacuerdo (2), totalmente en desacuerdo (1)	Competencia modelo CanMeds	Calificación global del profesorado
16	Es un ejemplo a seguir por su desempeño e integridad profesional	Profesional	0,86
17	Es un líder, es decir guía y ejerce influencia en el grupo con un objetivo	Líder	0,85
18	Respeto a los residentes	Profesional	0,88
19	Inspira confianza para plantearle problemas médicos, académicos o personales	Comunicador	0,84
20	Analiza y resuelve adecuadamente los conflictos	Lider	0,84
21	Es congruente entre lo que dice y hace	Profesional	0,85
22	Promueve un ambiente propicio para el aprendizaje	Académico	0,85
23	Refuerza la actitud de ayuda a los demás	Colaborador	0,84
24	Mantiene el orden de los procesos educativos de los residentes	Académico	0,85
Sección/ pregunta	3. Docencia en la clínica Opciones de respuesta: diario (4), alguna vez en la semana (3), alguna vez al mes (2), nunca (1), no aplica (0)	Competencia modelo CanMeds	Calificación global del profesorado
En la práctica el profesor interactúa con los residentes durante			
25	El paso de visita	Experto	0,84
26	Las sesiones de revisión de casos clínicos (expediente, estudios de gabinete y laboratorio, anatomopatológicos, morbimortalidad, etc.)	Experto	0,83
27	La consulta externa	Experto	0,82
28	La práctica en áreas diagnósticas y quirúrgicas	Experto	0,84

Sección/ pregunta	4. Docencia en el aula Opciones de respuesta: siempre (4), casi siempre (3), casi nunca (2), nunca (1), no aplica (0)	Competencia modelo CanMeds	Calificación global del profesorado
29	Asiste con regularidad a clase	Profesional	0,85
30	Presenta el programa al inicio del curso (objetivos, contenidos, actividades y criterios de evaluación)	Académico	0,86
31	Explica con claridad los temas de la asignatura	Académico	0,85
32	Emplea diversas estrategias didácticas para abordar los contenidos de la asignatura	Académico	0,83
33	Promueve el trabajo en equipo entre los residentes	Colaborador	0,84
34	Promueve el desarrollo de habilidades interpersonales, para una comunicación efectiva	Comunicador	0,84
35	Utiliza la estrategia de aprendizaje basado en problemas o casos para el en- tendimiento de los contenidos	Académico	0,85
36	Evalúa la comprensión y aplicación de los temas más que su memorización	Académico	0,87
37	Evalúa de manera continua el desempeño de los residentes	Académico	0,87
38	Atiende las dificultades del aprendizaje de los residentes	Académico	0,86

En la Tabla 2 se muestran los enunciados de los reactivos por secciones del cuestionario, y su clasificación de acuerdo con el modelo CanMeds, así como el promedio de calificación obtenida por el conjunto de los docentes evaluados (Fuente: elaboración propia).

Todos los ítems se encuentran en el rango de 0,88 a 0,79. Las preguntas mejor evaluadas (entre 0,88 y 0,86) se refieren al respeto a los residentes (profesionalismo), estar actualizado (experto) y ser un ejemplo a seguir (profesionalismo). En el aula se valora que el profesor presente el programa del curso al inicio, lleve a cabo una evaluación justa y continua, y atienda a los residentes con dificultades de aprendizaje, todos ellos forman parte de su rol como académico.

Se observó que las preguntas con los menores promedios (0,81 a 0,79) fueron de la 7 a la 10, todas ellas corresponden al rol académico y se refieren específicamente al uso de tecnología y recursos educativos digitales para la enseñanza y el aprendizaje, lo que constituye un área de oportunidad.

Por otro lado, se exploraron los casos de los profesores que obtuvieron una calificación global igual o mayor a 96. Se contabilizaron 1.377 cuestionarios, lo que corresponde a 137 profesores en total. De estos profesores, aquellos que alcanzaron un promedio de 98 o más sumaron 508 cuestionarios, equivalente a 57 profesores. Además, se identificaron 18 profesores evaluados con diez o más cuestionarios. Es importante dimensionar el tamaño de los grupos que son evaluados pues no todas las especialidades médicas cuentan con el mismo número de residentes.

Discusión

Los resultados del instrumento de ED del PUEM en el ciclo 2022-2023 revelan propiedades psicométricas destacadas (Downing *et al.*, 2003). Los resultados estadísticos demuestran una alta confiabilidad del instrumento, lo que cumple con el primer objetivo del estudio que es demostrar la validez y confiabilidad. Al igual que en el instrumento anterior, las pruebas de alpha de Cronbach y el análisis factorial fueron consistentes y estables. El análisis factorial confirmatorio corroboró la estructura del cuestionario en cuatro factores, mostrando que el profesionalismo está incorporado en las actividades académicas del profesorado. Gallé & Lingard (2010) y Vivas *et al.* (2021) subrayan la importancia del profesionalismo en el desempeño de los docentes, lo que se corrobora en la amplia población del PUEM.

El nuevo cuestionario aporta datos sobre el tiempo que los docentes dedican a diferentes actividades académicas. Se encontró que la mayoría del tiempo educativo se destina a la práctica clínica (48,35%), seguido por el aula (25,33%) y la tutoría (26,32%). Esto confirma que el aprendizaje en las especialidades médicas se logra principalmente en la interacción clínica (Domenjó, 2019).

Las calificaciones globales del profesorado por los residentes promediaron 0,85. Las puntuaciones más altas corresponden a profesionalismo, actividades académicas y docencia en el aula mostrando que las habilidades pedagógicas se expresan en diversas facetas. En contraste, las puntuaciones más bajas en el nuevo instrumento incluyen preguntas sobre el uso de tecnología educativa y comunicación a distancia, revelando un área de oportunidad para el profesorado del PUEM, que utiliza poco los recursos digitales en la enseñanza.

Se agregaron 10 preguntas en la sección 4 (preguntas 29 a 38), que indagan la dinámica educativa en el aula, incluyendo programa, contenidos, estrategias didácticas y evaluación. Las dos preguntas con porcentajes más elevados fueron la aplicación de los temas vistos en clase y la evaluación continua del aprendizaje de los residentes. Estos dos rasgos muestran la relación entre los temas de clase con la práctica clínica, así como, la supervisión de los docentes a los residentes por medio de la evaluación formativa (Graue & Argüero, 2015).

Las pruebas estadísticas realizadas para validar el constructo del instrumento se apegan a las reportadas en la literatura (Vázquez *et al.*, 2020). No obstante, al buscar literatura sobre la ED en residencias médicas, lo más cercano al ejercicio realizado fue la aplicación del PHEEM “*Postgraduate Hospital Educational Environment Measure*” (Roff *et al.* 2005) para medir el ambiente educativo en una sede del Instituto Mexicano del Seguro Social (IMSS) en Puebla, México (Hernández *et al.*, 2022), sin embargo, este cuestionario no es específico de las residencias médicas por lo que los datos no son comparativos (Martínez *et al.*, 2016).

La comparación de las cifras de validación es posible entre la versión anterior del instrumento y la actual. Los datos de validez y confiabilidad de la primera versión se han publicado en distintos momentos y escritos por ejemplo en el artículo de la revista Educación Médica (Hamui-Sutton *et al.*, 2017), en el capítulo 8 del libro: “Evaluación de las competencias docentes en las ciencias de la salud” (Hamui-Sutton *et al.*, 2019) y en el artículo publicado en 2023 donde se evaluaron las tendencias de las calificaciones en la ED de las residencias médicas longitudinalmente durante 8 años (Hamui-Sutton *et al.*, 2023).

Otro aspecto para considerar, además de las métricas y que también respalda los datos de la ED se refiere a la robustez del programa digitalizado con que opera la DEP. Olascoaga *et al.* (2021) describe algunas características necesarias para hacer funcionar un sistema de ED, estas son: a) diseño y operación digital por parte de la institución educativa que respalda el programa académico; b) naturaleza periódica y continua; c) uso de instrumentos de evaluación validados, confiables y de fácil aplicación; d) capacidad para procesar y organizar bases de datos; y e) generación de datos comprensibles para informar y realimentar a los docentes. En el caso de la ED del PUEM, todos estos aspectos se cumplen para asegurar que los procedimientos sean accesibles, eficientes, oportunos y justos.

A pesar de que se trataron de cuidar los aspectos del diseño, académicos y técnicos en el proceso de la ED, existen posibles sesgos derivados de la naturaleza subjetiva de las preguntas y la

dosis de afectividad en las respuestas de los residentes (Spooren & Mortelmans, 2006). Una de las restricciones del estudio se relaciona con la muestra, la diferencia entre aquellos con nombramiento universitario y los médicos en la clínica que fungen como profesores, pero no cuentan con una relación formalizada, los excluye de la población considerada. También se encontraron profesores titulares y/o adjuntos que cuentan con contrato, pero no atienden a los residentes en la práctica. Estas anomalías reducen el número de profesores evaluados y puede distorsionar la representación de la realidad.

Un aspecto adicional restrictivo es la escasez en ocasiones del número de cuestionarios por profesor debido a la dispersión de los residentes en diversas sedes y especialidades, lo que dificulta el análisis estadístico de los datos. Otra limitación del trabajo es la necesidad de estudios cualitativos como etnografías o entrevistas individuales y grupales para interpretar el contexto específico de cada sede y curso, permitiendo una comprensión más profunda de las interacciones y dinámicas en situaciones clínicas. Estos enfoques detallados cualitativos no pueden lograrse mediante encuestas a grandes poblaciones como las utilizadas en esta investigación (Hamui-Sutton, 2022).

En cuanto a futuras investigaciones, sería interesante explorar la influencia del desempeño docente en los resultados académicos y profesionales a largo plazo de los residentes. Se podría profundizar en las diferencias de percepción entre distintos grupos de residentes según el año de residencia o la especialidad médica (Vaižgėlienė *et al.*, 2017). Además, dado el avance rápido de la tecnología en la práctica médica, sería relevante investigar cómo integrar de manera efectiva nuevas tecnologías en la enseñanza clínica y evaluar su impacto en la formación de los residentes, especialmente considerando que el uso de tecnología educativa digital constituye un área de oportunidad entre los docentes. Como señala Cervantes *et al.* (2020), la mayoría de los residentes y profesores tienen acceso a dispositivos electrónicos y habilidades digitales, aunque actualmente se utilizan poco con fines educativos. Recursos como inteligencia artificial, simuladores, aplicaciones y bases de datos podrían potencialmente transformar la enseñanza y el aprendizaje en medicina si se implementan de manera efectiva.

Conclusiones

Este estudio resalta la importancia fundamental de la ED en las residencias médicas del PUEM. La validez y confiabilidad del cuestionario rediseñado ha sido rigurosamente validado, demostrando su capacidad para medir con precisión las competencias del profesorado en diferentes roles académicos. Los resultados muestran que los profesores son altamente valorados en aspectos

como el profesionalismo, la actualización de conocimientos y la evaluación continua del aprendizaje, así como el tiempo que dedican al aula, la clínica y la tutoría, revelando que el entrenamiento de los residentes se centra en la práctica. En estos contextos, los profesores son ejemplos que moldean el desarrollo profesional y el carácter de los residentes, idealmente de acuerdo con la ética médica y el humanismo. A su vez, revela áreas de mejora como el uso de tecnología educativa digital.

Estos hallazgos orientan estrategias institucionales para mejorar la calidad educativa. La implementación efectiva del proceso de la ED posibilita realimentar de manera individual a cada docente por medio de un reporte personalizado, y en lo social promueve un ambiente de aprendizaje continuo y colaborativo, crucial para el desarrollo profesional de los futuros especialistas médicos. En última instancia, la ED continua y basada en datos sólidos fortalece la calidad de la enseñanza en las especialidades médicas, apoyando así el desarrollo constante de los programas educativos y la satisfacción general de los residentes.

Reconocimientos

Agradecimientos: A Czeslaw Kristofer Lugowski Rivero y Elba Berrocal Pérez por su valioso apoyo en el manejo del sistema informático de la Evaluación Docente de la DEP y por organizar la base de datos que nos permitió realizar el análisis que aquí se presenta.

Fuentes de financiamiento: Ninguna

Declaración de conflictos de interés: Las autoras declaran no tener conflictos de interés.

Declaración de contribuciones: Las contribuciones de cada autor al trabajo fueron como sigue: LHS e IZG elaboraron el objeto de estudio, el diseño metodológico, la discusión de los resultados y la redacción del artículo, RFG realizó los análisis estadísticos, LHS, IZG y RFG trabajaron en la versión final del artículo.

Referencias

Anderson J & Gerbing, DW (1988). Modelado de ecuaciones estructurales en la práctica: una revisión y un enfoque recomendado de dos pasos. *Psychological Bulletin*, **103** (3), 411–423. DOI: 10.1037/0033-2909.103.3.411

Celis E, Dehesa E, Martínez A (2018). El residente como evaluador del desempeño docente en las especialidades médicas. *Educación Médica* **19**(4), 217-222. DOI: 10.1016/j.edumed.2017.03.021

Cervantes L, Janet M, Peña M, Alma A, Ramo A. (2020). Uso de las tecnologías de la información y comunicación como herramienta de apoyo en el aprendizaje de los estudiantes de medicina. *CienciaUAT* **15**(1), 162-171. DOI: 10.29059/cienciauat.v15i1.1380

Medina-Díaz M. del R, & Verdejo-Carrión A. (2020). Validez y confiabilidad en la evaluación del aprendizaje mediante las metodologías activas. *Alteridad* **15**(2), 270-284. DOI: 10.17163/alt.v15n2.2020.10

Domenjó, M. (2019). Aprendizaje y prácticas clínicas. *Educación Médica* **20**(2), 100-104. DOI: 10.1016/j.edumed.2018.12.019

Downing M. (2003). Validity: on meaningful interpretation of assessment data. *Medical education* **37**(9), 830–837. DOI: 10.1046/j.1365-2923.2003.01594.x

Frank J, Snell L, Sherbino J. (2015). CanMEDS 2015 Physician Competency Framework. Ottawa: Royal College of Physicians and Surgeons of Canada. Disponible en: <https://www.royalcollege.ca/content/dam/document/standards-and-accreditation/2015-can-meds-framework-reduced-e.pdf> consultado el 19 de noviembre de 2024

Gallé J & Lingard L. (2010). A medical student's perspective of participation in an interprofessional education placement: An autoethnography. *Journal of Interprofessional Care* **24**(6), 722–733. DOI: 10.3109/13561820903274954

Graue E & Argüero R. (2015) El aprendizaje observacional y el papel del docente como modulador de conductas. En *Educación Médica. Teoría y Práctica*, ed Sánchez Mendiola, *et.al.* pp. 283-288. Elsevier, Facultad de Medicina, UNAM.

Hamui-Sutton L, Fuentes R, Rodríguez S. (2023) Tendencias en las calificaciones de la evaluación docente en las residencias médicas. Un estudio longitudinal de ocho años. *Revista Iberoamericana de Evaluación Educativa* **16**(2), 39-54. DOI: 10.15366/riee2023.16.2.003

Hamui-Sutton A, Ortiz A, Cejudo L, Lavallo C, Vilar P. (2017). La evaluación de los docentes desde la perspectiva de los médicos residentes del Plan Único de Especializaciones Médicas. *Educación Médica* **18**(2), 89-97. DOI: 10.1016/j.edumed.2016.06.010

Hamui-Sutton L, Ortiz A, Gatica F. (2019). *Evaluación de las competencias docentes en las ciencias de la salud*. Editores Facultad de Medicina, Manual Moderno. UNAM. México. ISBN electrónico 9786073013918

- Hamui-Sutton L. (2022) *Interacciones narrativas en la clínica: Más allá del cerebro*. Libros Facmed, UNAM. México <https://libros.facmed.unam.mx/wp-content/uploads/2022/08/Libro-Interacciones-y-narrativas-en-la-clinica1082022.pdf>
- Hamui-Sutton L, Yáñez J, Rodríguez-Hernández MF (2024). La evaluación docente en las residencias médicas: ¿motivación o molestia para los profesores? *Fundación Educación Médica* **27**(2), 89-99. DOI: 10.33588/fem.272.1331.
- Hernández J, Ramírez L, Roco A, Fernández M. (2022). Ambiente clínico académico de las residencias médicas en el Instituto Mexicano del Seguro Social, Puebla. *Revista de Educación y Desarrollo* **62**, 37-64 https://www.cucs.udg.mx/revistas/edu_desarrolloanteriores/62/62_Hernandez.pdf
- Luchsinger L, Berthold A, Maeder M B, Giger M, Bauer W, Siegrist M. (2023). A questionnaire for quality control in postgraduate medical education in Switzerland. *Medical Teacher* **45**(9), 1012-1018. DOI: 10.1080/0142159x.2023.2168181
- Martínez A, Gómez F, Hernández I, Flores F, Sánchez M. (2016). Evaluación del desempeño docente en la especialización de Medicina Familiar. *Revista Médica del Instituto Mexicano del Seguro Social* **54**(5), 612-619.
- Numeralia-DEP (2024). *Numeralia Especializaciones Médicas PUEM. Ciclo 2022-2023*. Disponible en: <https://www.fmposgrado.unam.mx/index.php/numeralia-especializaciones-medicas-2024> consultado el 18 de noviembre de 2024.
- Olascoaga A, Aguirre R, López J. (2021). Evaluación de docentes en programas de posgrado. *ARS Medica, Revista de Ciencias Médicas* **46**(4), 44-50. DOI: 10.11565/arsmed.v46i4.1846
- Pritchard J, Alavian S, Soogoor A, Bartels S, Hall A. (2023). Global health competencies in postgraduate medical education: a scoping review and mapping to the CanMEDS physician competency framework. *Canadian medical education journal* **14**(1), 70-79. DOI:10.36834/cmej.75275
- Roff S, McAleer S, Skinner A. (2005). Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. *Medical teacher* **27**(4), 326-331. DOI: 10.1080/01421590500150874
- Spooren P, & Mortelmans D. (2006). Teacher professionalism and student evaluation of teaching: Will better teachers receive higher ratings and will better students give higher ratings? *Educational Studies* **32**(2), 201-214.
- Vaižgėlienė, E., Padaiga, Ž., Rastenytė, D., Tamelis, A., Petrikonis, K., & Fluit, C. (2017). Evaluation of clinical teaching quality in competency-based residency training in Lithuania. *Medicina* **53**(5), 339-347. DOI: 10.1016/j.medic.2017.08.002
- Vázquez M, Villar N, Fernández R, Cortés C. (2020). Validación e implementación de una metodología para la autoevaluación en la carrera de Medicina. *Medisur, revista electronica* **18**(4), 583-590. <https://www.redalyc.org/journal/1800/180064571008/html/>
- Vivas D, Reinoso N, Jaimes A. (2021). Profesionalismo médico como competencia, una visión desde la narrativa: estado del arte. *Educación Médica* **22**, S517-S520. DOI: 10.1016/j.edumed.2021.01.010