



**XVI**  
Congreso Nacional de  
Investigación Educativa  
CNIE-2021

## Un método para evaluar y mejorar la calidad de un caso de Aprendizaje Basado en Problemas

### **Kevin David Laguna Maldonado**

Jefaturad de la Unidad Departamental de Planeación Educativa, Universidad de la Salud  
[Kevin.lagunam@unisa.cdmx.gob.mx](mailto:Kevin.lagunam@unisa.cdmx.gob.mx)

### **Deyamira Matuz Mares**

Departamento de Bioquímica. Facultad de Medicina UNAM  
[deyabq@comunidad.unam.mx](mailto:deyabq@comunidad.unam.mx)

### **Teresa I. Fortoul Van der Goes**

Departamento de Biología Celular y Tisular. Facultad de Medicina UNAM  
[Fortoul@unam.mx](mailto:Fortoul@unam.mx)

Área temática 12. Evaluación educativa.

Línea temática: Evaluación y calidad educativa.

Tipo de ponencia: Reporte final de investigación.



### **Resumen**

**Objetivos:** el presente trabajo tiene como objetivo proponer un método que permita evaluar la calidad de un caso de ABP diseñado para un curso. Se ejemplificará el método a través de la evaluación de la calidad de un caso clínico para su aplicación en un curso de Bioquímica y Biología Molecular.

**Método:** se realizó un estudio de desarrollo, descriptivo típico, transversal, con un enfoque cuantitativo, para desarrollar un método que permita la evaluación de la calidad de un problema de ABP de manera sistemática. El método evalúa los casos en dos momentos por tres grupos de expertos (30 en total), así como por un grupo piloto de 7 estudiantes que recién terminaron el curso de Bioquímica y Biología Molecular. La evaluación se realizó con un instrumento que utiliza una escala tipo Likert. Para analizar los datos se calcularon medias aritméticas para los resultados de cada evaluación.

**Resultados:** al comparar ambas evaluaciones hechas por los expertos se observa una mejora estadísticamente significativa ( $W=257$ ,  $p<0.05$ ) lo que sugiere que el caso mejora después del uso del método propuesto. Asimismo, se calculó un alfa de Cronbach de entre 0.94 y 0.95 que demuestra una buena confiabilidad.

**Conclusiones:** el método propuesto permite sistematizar la evaluación de la calidad de los casos clínicos de ABP, recopilando en el proceso suficientes evidencias de validez que apoyen el resultado. Además, los resultados sugieren que dicho método mejora la calidad del caso.

**Palabras clave:** Aprendizaje Basado en Problemas, Calidad, Evaluación, Mejora, Caso.

## Introducción

En la actualidad se cuenta con una gran gama de posibles alternativas a la enseñanza tradicional, que estimulan el desarrollo de las competencias propias de una profesión, más allá de la adquisición de conocimiento o habilidades aisladas. Una de las estrategias más utilizadas es el Aprendizaje Basado en Problemas (ABP), el cual, en el contexto de la educación médica, tiene como fundamento la exposición de problemas clínicos similares a los que enfrentarán los estudiantes durante la práctica profesional, con el fin de desarrollar habilidades relacionadas con el diagnóstico, la elección del tratamiento y el análisis de la correlación básico - clínica (Barrows, 1996; Gurpinar, Kulac, Tetik, Akdogan, & Mamakli, 2013; Thistlethwaite et al., 2012).

En el desarrollo del ABP siempre intervienen tres actores fundamentales: el problema, el tutor y el estudiante (Gijsselaers & Schmidt, 1990). Aunque los tres factores tienen importancia en el desempeño de la estrategia, el problema tiene un papel primordial, ya que es el centro de la discusión y aprendizaje de los estudiantes y el punto de partida del tutor para guiar a los estudiantes en la resolución de éste (Keshk, El-Azim, & Qalawa, 2016).

El caso clínico debe seguir ciertas recomendaciones para ser un material adecuado en el ABP, por ejemplo: 1) debe ser parecido a la realidad, 2) debe ser adecuado para el nivel de conocimiento de los estudiantes, 3) debe permitir que surjan dudas para estimular la discusión e incitar a su vez la búsqueda de información, entre otras características. Algunos autores mencionan que si un caso cumple con estas características es porque tiene suficiente calidad (Niemer, Pfenndt, & Gers, 2010; Qin, Wang, & Floden, 2016; Valdés, 2006; Walker & Leary, 2009).

En 1990 Gijsselaers y Schmidt plantearon por primera vez la relación entre la calidad de los problemas de ABP y la efectividad de la estrategia (Gijsselaers & Schmidt, 1990), se analizaron variables como el conocimiento previo de los estudiantes, la correcta guía por parte del tutor, el buen trabajo dentro del grupo, el tiempo invertido por el estudiante en las sesiones individuales y la calidad de los problemas, siendo esta última el factor que más repercutía en la efectividad del ABP (Sockalingam, Rotgans, & Schmidt, 2012, 2011).

Desde entonces, diversos autores han indicado la importancia de la evaluación de la calidad de los problemas de ABP para asegurar que los estudiantes lleguen a los objetivos planteados y los tutores puedan guiar de forma correcta a los estudiantes a través del mar de información que supone el uso de esta estrategia, y con esto de asegurar al máximo los beneficios del ABP como estrategia didáctica (Keshk et al., 2016; Sockalingam et al., 2012).

Aún con la importancia que conlleva la calidad en los casos, no existe un método sistemático para evaluarla o mejorarla, por lo que el presente trabajo tiene como objetivo proponer un método sistemático que permita evaluar y mejorar la calidad de un caso clínico de ABP.

## Desarrollo

Se realizó un estudio de desarrollo, descriptivo típico, transversal, con un enfoque cuantitativo, que tenía como fin desarrollar un método que permitiera evaluar la calidad de un problema de ABP de manera sistemática. Es importante recalcar que dada la ambigüedad del concepto de calidad de los casos clínicos de ABP, en el presente trabajo se utilizó la definición de calidad propuesta por García y Martínez en el 2018, quienes definieron la calidad de un problema de ABP con base en el grado en que el problema tenga los siguientes seis factores; algunas de las definiciones presentadas se tomaron directamente del texto original (C. García Rivera, González, & Adrián., 2019; R. C. García Rivera, 2018):

- 1. Estimula el aprendizaje autodirigido.** Este factor se refiere al proceso en el que los alumnos asumen la iniciativa de su aprendizaje, realizan el diagnóstico de sus necesidades, formulan sus objetivos, identifican los recursos necesarios para aprender, eligen y aplican estrategias de aprendizaje y evalúan los resultados obtenidos.
- 2. Coincide con el nivel previo de conocimientos de los estudiantes.**
- 3. Promueve la toma de decisiones.** Se refiere al grado en que los estudiantes, con base en los conocimientos previos, hechos e información lógica y fundamentada, pueden decidir sobre el diagnóstico, tratamiento, pronóstico, etc., en el problema que se les presenta.
- 4. Estimula el pensamiento, análisis y razonamiento.** Este factor se relaciona con el grado en que el problema permite al estudiante desarrollar actividades cognitivas para lograr una comprensión más profunda de los temas discutidos.
- 5. Aumenta el interés en el tema.** Se espera que el problema de ABP sea atractivo e interesante para lograr captar la atención de los estudiantes.
- 6. Formato.** En este factor se explora que el problema se presente en el formato y longitud adecuados, y con suficiente claridad, incluyendo el uso de imágenes o vídeos.

Es importante mencionar que dicha definición de calidad en un caso de ABP retoma muchos de los planteamientos hechos por otros autores. Aunado a esto, el trabajo de García y Martínez, es de los pocos trabajos en el que se presenta una herramienta, con una descripción fehaciente de las evidencias de validez, que permite evaluar la calidad en los problemas.

A continuación, se resumen los pasos del método para evaluar la calidad de los problemas de ABP:

### 1. Elaboración de los casos clínicos

Como primer paso, en este trabajo se desarrollaron cuatro casos clínicos de ABP. Para dicho fin se creó un grupo de cuatro profesores que siguieron las recomendaciones de Azer en el 2012 para la elaboración de casos

clínicos de tipo ABP. Cada caso estaba conformado por tres partes: una primera en la que se presentaba al paciente y su padecimiento, otra en donde se analizaban los estudios de laboratorio y gabinete y una última en la que se daba un cierre al caso.

Con el fin de mostrar el método propuesto en el presente trabajo se analizará sólo uno de los casos, el primero que corresponde a un paciente con deshidratación.

## 2. Grupos de expertos

Con el fin de evaluar la calidad desde la perspectiva de todos los ámbitos posibles y de contrastar la visión de diferentes grupos de expertos, se propuso que tres grupos diferentes evaluaran el mismo caso. Para crear cada grupo se utilizó un muestreo no probabilístico por conveniencia. Dichos grupos estuvieron formados de la siguiente manera:

- 15 expertos en Bioquímica y Biología Molecular, que evaluaron el contenido correspondiente a los conocimientos de Bioquímica y Biología Molecular básicos en el ejercicio de la Medicina
- 15 expertos en Medicina, que evaluaron el contenido correspondiente a los signos, síntomas y estudios de laboratorio y gabinete expuesto en el problema
- 15 expertos en ABP, que evaluaron el formato, la extensión y las herramientas que se incluyen en los problemas.

Los criterios de inclusión, exclusión y eliminación que se utilizaron para cada grupo fueron:

- Criterios de inclusión:
  - o Ser participante activo en el área de expertiz, es decir, ser profesor, creador de material de ABP o médico en ejercicio profesional.
  - o Tener al menos 5 años de experiencia en el área correspondiente.
  - o Aceptar participar como experto evaluador.
  - o No haber participado en la creación de los problemas de ABP.
- Criterios de exclusión:
  - o No estar ejerciendo en el área de expertiz actualmente.
  - o Tener menos de 5 años de experiencia.
  - o Haber participado en la creación de los problemas de ABP.

- Criterios de eliminación:
  - o Expresar la voluntad de abandonar el proyecto.
  - o No entregar las evaluaciones solicitadas en tiempo.
  - o No entregar la evaluación de alguno de los casos.

### 3. Primera evaluación

Una vez formados los grupos de expertos se sometió a evaluación, por primera vez, la calidad de los problemas con la herramienta propuesta por García y Martínez en el 2018. La herramienta de evaluación es un cuestionario que constó de 20 reactivos que se responde a través de una escala tipo Likert con cinco opciones de respuesta (C. García Rivera et al., 2019; R. C. García Rivera, 2018).

A través del uso de dicha herramienta se buscó establecer el grado inicial en que el problema de ABP cumplía con cada uno de los factores descritos en la definición de calidad (Downing, 2003).

Para que cada uno de los expertos respondieran el cuestionario se envió a través del correo electrónico los reactivos utilizando la herramienta "Google forms". Además del cuestionario, se envió a cada experto el caso clínico a evaluar y la lista de los objetivos anuales del curso donde se impartirían.

### 4. Grupos pilotos

Con base en los comentarios de los expertos se realizaron los cambios pertinentes a cada uno de los problemas de ABP. Posteriormente se convocó a un grupo para un estudio piloto, que podía estar integrado por entre 5 a 10 estudiantes que recién terminaban la asignatura de Bioquímica y Biología Molecular de la licenciatura de médico cirujano de la Facultad de Medicina de la UNAM. Para dicho fin se utilizó un muestreo no probabilístico por conveniencia.

La sesión con el grupo piloto duró alrededor de 3 horas. En dicho tiempo se utilizó el método de ABP para presentar las tres partes del caso clínico con un tutor guiando la actividad. Cuando se terminaba el abordaje de cada una de las partes se daba alrededor de 20 minutos para que los estudiantes buscaran la información necesaria para resolver los objetivos de aprendizaje planteados en la discusión, por lo que los estudiantes contaban con acceso a la web y/o con los libros para dicho fin. Una vez que se discutía la información obtenida, con la guía del tutor, se planteó la siguiente parte del caso repitiendo el método hasta finalizar todas las partes.

Es importante mencionar que, durante la sesión con el grupo piloto, se contó con la participación de dos observadores externos a la actividad, cuya función únicamente fue observar y evaluar el uso de la herramienta.

Las opiniones de los estudiantes se recopilaron nuevamente a través del instrumento de García y Martínez, 2018 en el mismo formato que se utilizó con los expertos y posteriormente se analizaron los resultados con los mismos estadísticos planteados para la evaluación por los expertos.

## 5. Segunda evaluación

Con los resultados del grupo piloto se realizaron los últimos cambios en el problema. Esta versión fue enviada nuevamente a los grupos de expertos que respondieron la primera vez con el fin de obtener una nueva evaluación de la calidad que se pudiera comparar con la primera.

### Marco conceptual

En el presente trabajo utilizamos el marco conceptual de validez propuesto por Messick (Messick, 1987). En él se reconocen cinco fuentes de evidencia de validez que apoyan los análisis realizados a los resultados de una evaluación (Carrillo, Sánchez, & Leenen, 2020). En este trabajo se busca obtener la evidencia de validez del contenido y de la estructura interna.

### Análisis estadístico

Para analizar y comparar los resultados se calculó la frecuencia de las opciones de respuesta para cada uno de los ítems (ver más adelante). Una vez ordenadas las variables se obtuvo la media de respuesta para cada uno de los factores y la media general del caso, y se asignó un nivel de calidad de acuerdo con los criterios reportados por García y Martínez 2018 (Tabla 1).

Las pruebas estadísticas mostraron que los datos no cumplían con los requisitos para realizar estadística paramétrica, por lo que se analizaron los datos con las siguientes pruebas:

- Se calculó el estadístico kappa de Fleiss para evaluar la concordancia entre los evaluadores al responder.
- Se calculó el alfa de Cronbach para evaluar la confiabilidad en la respuesta.
- Se realizó una prueba de Wilcoxon para analizar si existía una diferencia estadísticamente significativa entre la primera y la segunda evaluación de la calidad para cada uno de los problemas.

Todas las respuestas de las evaluaciones se analizaron con el programa estadístico PRISM 8, San Diego, California, Estados Unidos de Norteamérica (Licencia GPS-1694331-TCRF-288C0).

### Componente ético

El presente proyecto está registrado y aprobado por las Comisiones de Investigación y Ética de la División de Investigación de la Facultad de Medicina de la UNAM, y fue aceptado el 5 de febrero de 2019 con el número de oficio FM/DI/118/2018.

Además, se solicitó el consentimiento libre e informado de todos los participantes de acuerdo con la Ley General de Salud Artículo 21. Los resultados del estudio se publican de manera anónima y confidencial. No hubo repercusiones, positivas o negativas, para los participantes.

## Resultados

Para evaluar la calidad de los problemas de ABP, se sometieron a revisión, en una primera ocasión, cada uno de los problemas por los expertos, usando los criterios descritos anteriormente. Inicialmente, la muestra de expertos invitados era de 45 en total, 15 expertos por cada grupo. Sin embargo, únicamente respondieron 30 (Tabla 2), por lo que se obtuvieron 30 respuestas a través del cuestionario en línea, uno para cada uno de los problemas.

Como producto de esa primera evaluación se obtuvo la opinión de los expertos con respecto de la calidad inicial de los problemas de ABP. Posteriormente se realizó la aplicación de la herramienta en cuatro grupos piloto, uno para cada caso, en los que se pidió a dos observadores que analizaran el funcionamiento de la herramienta mientras se desarrollaba la sesión. Una vez terminada la sesión se solicitó que tanto los observadores como los estudiantes evaluaran el desempeño de los casos a través de la herramienta digital antes descrita. En esta evaluación se obtuvieron siete respuestas.

Posterior a la prueba en los grupos piloto, y tras hacer los cambios correspondientes a los problemas, se solicitó una segunda evaluación a los 30 expertos. Para esta segunda vez, se obtuvieron 27 respuestas en total, se perdió un experto del grupo clínico y dos del grupo de Bioquímica.

Para evidenciar si hubo cambios estadísticamente significativos en la calidad del problema 1 entre la primera y la segunda evaluación por parte de los expertos, se calculó la media de respuesta general para las dos evaluaciones (Figura 1). Se puede ver que según lo reportado en la tabla 1, el caso clínico en la primera evaluación se encontraba en el límite bajo del nivel de calidad “bueno” y tras la prueba con el grupo piloto y la realización de los cambios pasó a estar, según los expertos, en el límite alto de la misma categoría ( $W=257, p<0.05$ ).

En el caso de la evaluación hecha por los estudiantes y observadores durante la aplicación del grupo piloto 1, la media de la frecuencia de respuesta se reporta en la figura 2.

## Conclusiones

Como se planteó al inicio, el objetivo del presente trabajo era desarrollar un método que permitiera evaluar y mejorar la calidad de un problema de ABP, idealmente antes de ser usado en un curso. Al respecto, Gijsselaers

y Schmidt plantean que la calidad del material utilizado en el problema de ABP es el factor que más correlaciona con el aumento del interés en el tema, un aumento del tiempo que se invierte en las sesiones individuales, una mejora en la discusión y en el trabajo en equipo y facilita el trabajo de los tutores al guiar una discusión sobre el caso (23). Por tanto, evaluar la calidad de los problemas de ABP es imprescindible para asegurar la adecuada implementación de dicha estrategia.

Para el caso clínico evaluado por este método, los resultados muestran que cuenta con una calidad entre “Buena” y “Excelente”, según las evaluaciones de los expertos y de los grupos piloto. Aunado a lo anterior, la prueba de Wilcoxon permitió identificar una diferencia estadísticamente significativa entre la primera y la segunda evaluación de los casos por los expertos. Por tanto, podemos inferir que el método que se plantea permite evaluar y al mismo tiempo mejorar la calidad de los problemas de ABP.

Con respecto a la evaluación, es importante recalcar que con tres diferentes grupos de expertos se puede contrastar las opiniones de diversos campos de la medicina y la educación médica, y no limitarnos al evaluar la calidad desde cada perspectiva.

En el presente trabajo también se propuso obtener dos de las fuentes propuestas en el marco conceptual de Messick, la evidencia de validez del contenido y de la estructura interna (Messick, 1987).

Para obtener la evidencia de validez del contenido se creó un grupo de trabajo para diseñar los casos clínicos y se usó como referencia las causas de hospitalización, enfermedad y muerte más comunes en México, así como las guías de práctica clínica mexicanas e internacionales más actuales (INEGI, 2018, 2019; Soto-Estrada, Moreno-Altamirano, & Pahua, 2016). Una vez creados se aumentó la evidencia de validez al solicitar que tres grupos de expertos en diferentes áreas evaluaran los casos desde su zona de experiencia. Con dicho método se buscaba obtener opiniones diferentes que permitieran asegurar la mejor evaluación de la calidad en el caso.

Por último, con el fin de analizar la concordancia entre los expertos y fortalecer la evidencia de validez del contenido, se calculó el coeficiente kappa. Sin embargo, para la mayoría de las evaluaciones se obtuvo una kappa que describe un nulo acuerdo entre los expertos (Tang, Hu, Zhang, Wu, & He, 2015; Westphal et al., 2014). No obstante, hay que recordar que dicho estadístico depende de las 5 opciones de respuesta, y dado que casi la mitad de los expertos contesta “completamente de acuerdo” y la otra mitad contesta “de acuerdo”, el análisis estadístico mostrará que no hay acuerdo entre los expertos. Sin embargo, de forma empírica se puede decir que la mayoría de los evaluadores encuentra favorables los casos al evaluarlos, en especial la segunda vez.

En el caso de la evidencia de validez de la estructura interna se calculó la calidad de los casos (González, 2019). Dicho análisis muestra que los problemas de ABP tienen una suficiente evidencia de validez de la estructura con relación al constructo evaluado. Asimismo, para complementar esta evidencia se calculó el alfa de Cronbach para todas las evaluaciones realizadas (Cervantes, 2005; Taber, 2018). Según lo reporta la literatura, la mejor forma de obtener e interpretar dicho estadístico es a través del análisis de una condición (por ejemplo, problema

1 y primera evaluación por los expertos, tabla 3), y no calculando un valor que incluya a todas las condiciones (12 condiciones, tabla 3), lo cual resultaría confuso y ocultaría los valores individuales.

En las evaluaciones por parte de los expertos se obtiene una alfa de Cronbach alta, lo cual indica una alta confiabilidad de la evaluación y nos permite asegurar que muy probablemente, si se repitiera la evaluación, se obtendrían datos similares. En el caso del grupo piloto se obtuvo una alfa de Cronbach más baja; suponemos que este efecto se debe a los pocos participantes que hubo en dichos grupos.

Con base en todo lo anterior, se puede concluir que el método presentado en este estudio permite evaluar y mejorar la calidad de un caso clínico de ABP de manera válida y confiable. Sin embargo, es importante mencionar algunas limitaciones con respecto a los resultados obtenidos, por ejemplo:

- Al usar a los mismos expertos en ambas evaluaciones (método de tipo pre – postest) se limita el efecto que tienen las mejoras en los casos clínicos, ya que existe un sesgo por haber visto la primera versión del caso. En el presente proyecto se usó dicho método por la dificultad en conseguir a los expertos.
- Aunado a lo anterior, se usó la misma herramienta en ambas evaluaciones por los expertos, lo que supone una amenaza a la validez de la evaluación, por varianza irrelevante del constructo (Downing & Yudkowsky, 2009).

## Tablas y figuras

Tabla 1. Niveles de calidad de los problemas de ABP

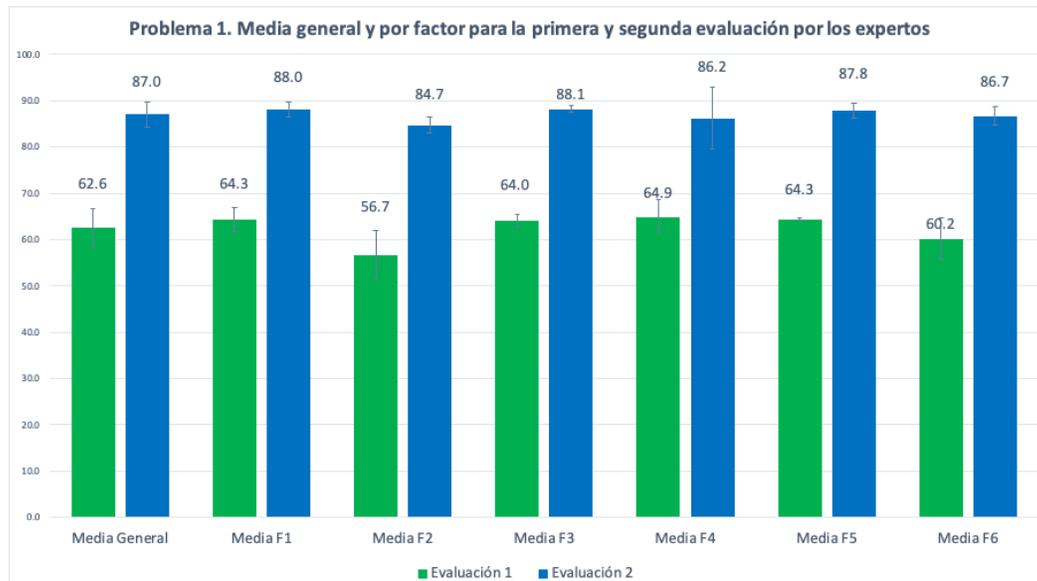
Nivel	Intervalo	
Deficiente	20	44
Regular	45	70
Buena	71	87
Excelente	88	100

Tabla 2. Expertos en la primera evaluación

Grupo de expertos	Sexo		Número de expertos total que respondieron por grupo	Porcentaje del total
	Masculino	Femenino		
ABP	4	3	7	23.3%
Bioquímica	8	5	13	43.3%
Clínica	8	2	10	33.31%
Totales	20	10	30	100%

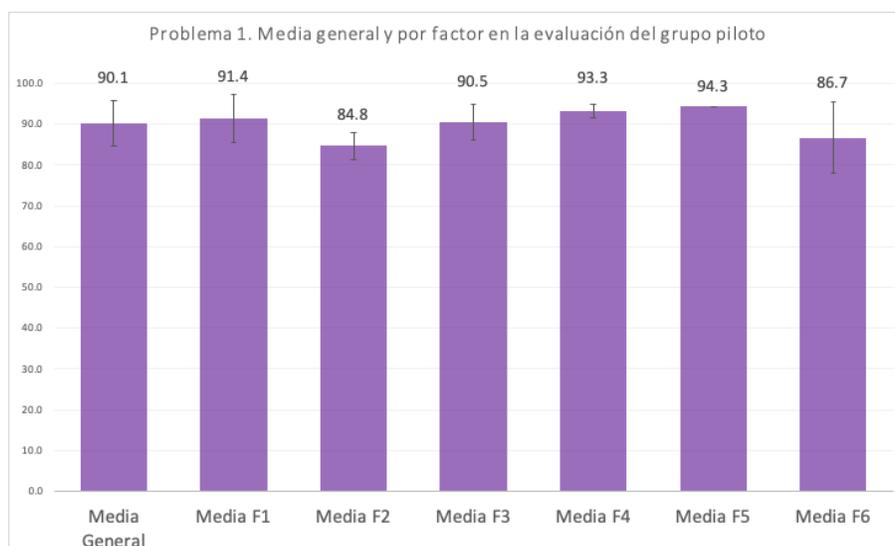
Se muestran los números brutos y porcentajes de los expertos en cada grupo que respondió la primera evaluación.

Figura 1. Diferencia de medias general y por factor en la primera y segunda evaluación por los expertos para el problema 1



Se presenta la diferencia entre las medias generales de evaluación y para cada uno de los componentes que evaluaron los al primer grupo piloto del primer problema. Las barras representan desviación estándar.

Figura 2. Diferencia de medias general y por factor en el grupo piloto para el problema 1



Las barras representan desviación estándar.

## Referencias

- Alexandre, J. (n.d.). When is statistical significance not significant?, (1), 31–55.
- Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning*, 1996(68), 3–12. <https://doi.org/10.1002/tl.37219966804>
- Carrillo, B. A., Sánchez, M., & Leenen, I. (2020). El concepto moderno de validez y su uso en educación médica. *Revista de Investigación En Educación Médica*, 9(33), 98–106.
- Cervantes, V. H. (2005). Interpretaciones Del Coeficiente Alpha De Cronbach. *Avances En Medición*, 3(December), 9–28.
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, 37(9), 830–837. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>
- Downing, S. M., & Yudkowsky, R. (2009). *Assessment in Health Professions Education*. (S. M. Downing & R. Yudkowsky, Eds.) (Second Edi). New York, NY.
- García Rivera, C., González, M., & Adrián. (2019). Calidad de los problemas de ABP. Evidencia de validez de un instrumento. *Revista de Investigación En Educación Médica*, 8(29), 58–68.
- García Rivera, R. C. (2018). *Calidad de los casos de Aprendizaje Basado en Problemas utilizados en la asignatura de Integración Básico - Clínica*. Universidad Nacional Autónoma de México.
- Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. *Statistical Science*, 24(3), 343–360. <https://doi.org/10.1214/09-STS301>
- Gijselaers, W., & Schmidt, H. (1990). Development and evaluation of a causal model of problem-based learning. In S. P. Co. (Ed.), *Innovation in Medical Education: An Evaluation of Its Present Status*. (pp. 95–113). New York: Springer Publishing Co.
- González, A. (2019). *Evidencias de validez en la elaboración y aplicación de casos de aprendizaje basado en problemas (ABP) para la asignatura de informática biomédica de la Facultad de Medicina de la UNAM*. Universidad Nacional Autónoma de México.
- Gurpinar, E., Kulac, E., Tetik, C., Akdogan, I., & Mamakli, S. (2013). Do learning approaches of medical students affect their satisfaction with problem-based learning? *AJP: Advances in Physiology Education*, 37(1), 85–88. <https://doi.org/10.1152/advan.00119.2012>
- Hanusz, Z., Tarasinska, J., & Zielinski, W. (2016). Shapiro–Wilk test with known mean. *Revstat Statistical Journal*, 14(1), 89–100.
- INEGI. (2018). *Comunicado de prensa Núm. 525 / 18 31 de octubre de 2018 características de las defunciones registradas comunicación social*. Ciudad de México, México.
- INEGI. (2019). *Comunicado de prensa Núm. 472 / 19 30 de septiembre de 2019 estadística de salud en establecimientos particulares 2018 generada a partir de la información de 2 863 establecimientos de salud*. Ciudad de México, México.
- Keshk, L. I., El-Azim, S. A., & Qalawa, S. A. (2016). Quality of Problem Based Learning Scenarios at College of Nursing in Egypt and KSA: Comparative Study. *American Journal of Educational Research*, 4(9), 701–710. <https://doi.org/10.12691/education-4-9-10>
- Mario, B., Biomed, C., Res, T., & Cesana, B. M. (2018). What p-value must be used as the Statistical Significance Threshold ?  $P < 0.005$ ,  $P < 0.01$ ,  $P < 0.05$  or no value at all ?, 5(4), 5310–5318. <https://doi.org/10.26717/BJSTR.2018.06.00135>
- Messick, S. (1987). Validity. In *Educational Measurement* (3rd Editio, p. 213). New York.

- Nierner, L., Pfenndt, K., & Gers, M. (2010). Problem-based learning in nursing education: A process for scenario development. *Nurse Educator*, 35(2), 69–73. <https://doi.org/10.1097/NNE.0b013e3181ced891>
- Qin, Y., Wang, Y., & Floden, R. E. (2016). The Effect of Problem-Based Learning on Improvement of the Medical Educational Environment: A Systematic Review and Meta-Analysis. *Medical Principles and Practice*, 25(6), 525–532. <https://doi.org/10.1159/000449036>
- Sockalingam, N., Rotgans, J. I., & Schmidt, H. G. (2012). Assessing the Quality of Problems in Problem-Based Learning. *International Journal of Teaching and Learning in Higher Education*, 24(1), 43–51.
- Sockalingam, N., Rotgans, J., & Schmidt, H. G. (2011). Student and tutor perceptions on attributes of effective problems in problem-based learning. *Higher Education*, 62(1), 1–16. <https://doi.org/10.1007/s10734-010-9361-3>
- Soto-Estrada, G., Moreno-Altamirano, L., & Pahua, D. (2016). Panorama epidemiológico de México, principales causas de morbilidad y mortalidad. *Revista de La Facultad de Medicina*, 59(6), 8–22.
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tang, W., Hu, J., Zhang, H., Wu, P., & He, H. (2015). Kappa coefficient: a popular measure of rater agreement. *Shanghai Archives of Psychiatry*, 27(1), 62–67.
- Thistlethwaite, J. E., Davies, D., Ekeocha, S., Kidd, J. M., MacDougall, C., Matthews, P., ... Clay, D. (2012). The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23. *Medical Teacher*, 34(6), 142–159. <https://doi.org/10.3109/0142159X.2012.680939>
- Valdés, N. (2006). *Elaboración y validación de casos de Aprendizaje Basado en Problemas (ABP) para el programa de Biología en Educación Medica Superior de la UNAM.*
- Walker, A., & Leary, H. (2009). A Problem Based Learning Meta Analysis: Differences Across Problem Types, Implementation Types, Disciplines, and Assessment Levels. *Interdisciplinary Journal of Problem-Based Learning*, 3(1), 6–28. <https://doi.org/10.7771/1541-5015.1061>
- Westphal, M., Leahy, R. L., Pala, A. N., Wupperman, P., Wang, Q., Li, W., ... ALARCÓN, L. F. (2014). Validez De Contenido Y Juicio De Expertos: Una Aproximación a Su Utilización. *Eating Behaviors*, 27(3), 313–318. <https://doi.org/10.1016/j.eatbeh.2007.10.005>