



XVI
Congreso Nacional de
Investigación Educativa
CNIE-2021

Desarrollo de una herramienta para evaluar el pensamiento crítico en estudiantes de Medicina

Pardo Vázquez Juan Pablo

Departamento de Bioquímica, Facultad de Medicina, UNAM
pardov@bq.unam.mx

Matuz Mares Deyamira

Departamento de Bioquímica, Facultad de Medicina, UNAM
deyabq@comunidad.unam.mx

Briz Tena Oliva

Departamento de Bioquímica, Facultad de Medicina, UNAM
oli_bt4291@comunidad.unam.mx

Área temática 12. Evaluación educativa.

Línea temática: diseño y validación de instrumentos.

Tipo de ponencia: reportes parciales de investigación.



Resumen

Introducción: el pensamiento crítico promueve la autonomía racional, la libertad intelectual y la investigación basada en evidencias para la solución de problemas. La falta de éste se ha relacionado con el incremento en el error médico que en algunos países supone la tercera causa de muerte, por lo que es necesario tener evidencia que los estudiantes desarrollan esta competencia desde los primeros años de la licenciatura.

Objetivo: diseñar una herramienta para evaluar el pensamiento crítico en estudiantes de Medicina que cuente con evidencias de validez.

Diseño/Método: estudio de desarrollo, cuantitativo, descriptivo, transversal. Se diseñó una herramienta de evaluación utilizando el método de exámenes de razonamiento clínico basados en casos (CBCRT por sus siglas en inglés).

Participantes: profesores de Bioquímica y Biología Molecular, médicos generales y médicos familiares, urgenciólogos, internistas, nefrólogos y estudiantes del primer año de la licenciatura de médico cirujano.

Resultados: se elaboró una herramienta con 24 reactivos siguiendo el proceso para obtener evidencias de validez, se realizó la psicometría con base en la Teoría Clásica del Test. Se obtuvo un α de Cronbach de 0.43 atribuido a una muestra pequeña y por lo tanto una varianza elevada, ya que era la primera vez que los estudiantes respondían este tipo de herramienta. Se identificaron 7 reactivos que no cumplieron con los parámetros estadísticos establecidos por lo que se revisarán con los comités especialistas. Se encontró una buena aceptación por parte del estudiantado y el tiempo utilizado fue menor al que utilizan los estudiantes para responder exámenes tradicionales.

Palabras clave: pensamiento crítico, evaluación, validez.

Introducción

A finales del siglo XX, surge el proyecto Tunning, que presenta un nuevo modelo de enseñanza enfocado en competencias, con el propósito de estimular el desarrollo de conocimientos, habilidades y actitudes que le permitan al estudiante enfrentarse a contextos sociales cambiantes y saturados de información (Castañeda Rivas, 2015; Sánchez Mendiola, 2018).

De acuerdo con lo anterior, en el año 2010, la Facultad de Medicina de la UNAM propuso modificaciones al plan existente e implementó un currículo mixto, estructurado por asignaturas y guiado por las siguientes competencias (Sánchez Mendiola et al, 2011):

1. Pensamiento crítico, juicio clínico, toma de decisiones y manejo de información
2. Aprendizaje autorregulado y permanente
3. Comunicación efectiva
4. Conocimiento y aplicación de las ciencias biológicas, sociomédicas y clínicas en el ejercicio de la medicina
5. Habilidades clínicas de diagnóstico, pronóstico, tratamiento y rehabilitación
6. Profesionalismo, aspectos éticos y responsabilidades legales
7. Salud poblacional y sistema de salud: promoción de la salud y prevención de la enfermedad
8. Desarrollo y crecimiento personal

Una de las competencias de mayor importancia es la primera. Makary y Daniel reportan que la tercera causa de muerte en Estados Unidos es por el error médico (Makary y Daniel, 2016) y una de las razones podría ser la falta de pensamiento crítico, por lo que la carencia de esta capacidad reflexiva es un factor importante en la existencia de médicos incapaces de tomar decisiones clínicas efectivas, lo que conduce a incidentes adversos (Zayapragassarazan, Menon, Kar y Batmanabane, 2016). Por ello, este modelo educativo exige no sólo nuevas estrategias de enseñanza, sino distintas herramientas para evaluar las competencias y asegurar que los egresados de la licenciatura cuenten con ellas.

En 1987, la Asociación Filosófica Americana (APA por sus siglas en inglés, *American Philosophical Association*) define el pensamiento crítico como el *juicio útil y autorregulado que resulta en interpretación, análisis, evaluación e inferencia, así como la explicación de las consideraciones probatorias, conceptuales, metodológicas, de criterio o contextuales sobre las que este juicio está basado* (Facione, 1998).

Algunos autores consideran que el pensamiento crítico y el razonamiento clínico son dos constructos diferentes (Kuiper, 2004), mientras que otros refieren que el pensamiento crítico aplicado a una situación clínica es el razonamiento clínico en sí (Noll, Key y Jensen, 2001) y es la base para la toma de decisiones que afectan la

salud y la calidad de vida de los pacientes (Viesca Treviño, Ponce de León y Sánchez Mendiola, s.f.). Debido a que en este proyecto coincidimos con las reflexiones de Facione, quien considera que el pensamiento crítico es específico para cada área y depende de conocimiento previo, nos ceñimos a esta segunda definición. De ahí, que utilicemos un método de evaluación basado en casos clínicos como base para el diseño de la herramienta propuesta (ten Cate, 2018).

El pensamiento crítico en contextos particulares, como sería el de la asignatura de Bioquímica y Biología Molecular - incluida en el currículo de la licenciatura de médico cirujano, impartida en el primer año y organizada en cuatro bloques: agua, metabolismo de carbohidratos, metabolismo de lípidos y genética-, demanda conocimiento específico (Facione, 1998) y tiene un papel importante en la comprensión de muchos temas en medicina, especialmente los conceptos relacionados con el metabolismo (Sadi, 2013). Por lo anterior, es difícil que los estudiantes repliquen el razonamiento del experto, si no cuentan con un conocimiento estructurado y experiencia propia, meta que sólo se puede cumplir si el estudiante está en contacto con casos clínicos durante su formación (Viesca et al, s.f.).

Si bien los estudiantes de medicina deben retener grandes cantidades de información a lo largo de su formación, sus habilidades para la solución de problemas y para la toma de decisiones son deficientes (Fortoul Van der Goes et al, 2012). Podría decirse, entonces, que tienen un nivel bajo de pensamiento crítico (Fortoul Van der Goes et al, 2012; Zayapragassarazan, 2016), ya que no es suficiente “recibir ideas, almacenarlas, recuperarlas y comunicarlas, sino que, idealmente, debe haber un proceso en el cual uno piensa, formula preguntas, busca y encuentra información relevante y llega a sus propias conclusiones” (Carrasco, 2018).

De acuerdo con el plan de estudios 2010 (Sánchez Mendiola et al, 2011), los atributos del pensamiento crítico que se espera que el estudiante de primer año logre, son:

- Identificar los elementos que integran el método científico y las diferencias para su aplicación en las áreas biomédica, clínica y sociomédica.
- Identificar, seleccionar, recuperar e interpretar, de manera crítica y reflexiva, los conocimientos provenientes de diversas fuentes de información para el planteamiento de problemas y posibles soluciones.
- Demostrar la capacidad para analizar, discernir y disentir la información en diferentes tareas para desarrollar el pensamiento crítico.

Si bien ya existen herramientas para evaluar el pensamiento crítico como el Watson-Glaser Critical-thinking Appraisal (WGCTA) y el California Critical thinking Disposition Inventory (CCTDI), estos evalúan dicho fenómeno de manera no específica (Paul, 2014).

En este proyecto se planteó el desarrollo de un instrumento de medición del pensamiento crítico ya que a pesar de que en el departamento de Bioquímica se han hecho esfuerzos por trabajar con casos clínicos, no contamos

con herramientas que tengan suficientes evidencias de validez. El instrumento propuesto se enfocó en evaluar únicamente los dos últimos atributos planteados por la Facultad de Medicina. Estos coinciden con los rubros de interpretación y análisis de la definición de Facione que incluyen actividades como clarificar, examinar ideas y detectar y analizar argumentos (Facione, 1998).

Desarrollo

Método

Se realizó un estudio de desarrollo, cuantitativo, descriptivo, transversal en el Departamento de Bioquímica de la Facultad de Medicina de la UNAM.

Con base en el método de Downing y Haladyna (Lane, Raymond, Haladyna y Downing, 2016) se realizaron los primeros pasos para el desarrollo de la herramienta. Se convocó a un grupo de profesores y/o médicos con cinco años o más de experiencia en docencia en la asignatura de Bioquímica y Biología Molecular, en la elaboración de tablas de especificaciones y en reactivos, con lo que se conformó un consejo rector (CR) con tres médicos cirujanos y una bióloga; un comité de diseño (CD) con dos médicos cirujanos, dos biólogos y un químico farmacéutico biólogo; un comité de elaboración de reactivos (CER), con dos médicos cirujanos, cinco biólogos y un químico farmacéutico biólogo y dos comités de validación (CV), el primero (CV1) conformado por tres médicos cirujanos, dos biólogos y un químico farmacéutico biólogo y el segundo (CV2) conformado por dos biólogos, un químico farmacéutico biólogo, tres médicos cirujanos, dos médicos internistas, un médico oncólogo, siete médicos residentes de medicina familiar, un médico residente de nefrología y tres médicos residentes de medicina interna.

Con base en el programa de la asignatura (Facultad de Medicina, 2018), el nivel académico de los estudiantes y el objetivo de la evaluación, el CR determinó que el escenario clínico ideal para probar la herramienta era diabetes tipo 2 (DT2) por lo que se tomaron en cuenta los tres primeros bloques de la asignatura. Con estos lineamientos, el CD seleccionó 66 de 162 objetivos de aprendizaje que consideró relevantes para esta patología y propuso los resultados de aprendizaje -150-. Estos se dividieron en ocho partes, se enviaron a los miembros del CER y se les solicitó elaborar dos reactivos de opción múltiple (con 3 opciones de respuesta) por cada resultado de aprendizaje para conformar un examen de conocimientos generales con la intención de contar con una herramienta para obtener evidencias de validez de relación con otras variables. Una vez que se recibieron los reactivos, se convocó a una reunión con el CD, se seleccionaron 44 y se enviaron al CV1 junto con un cuestionario (Rivera, Flores, Alpuche y Martínez, 2016) para calificar la estructura de los mismos y con base en los resultados se convocó a una junta en la que se hicieron las modificaciones necesarias para la conformación final del instrumento.

En función de los resultados de aprendizaje, los reactivos y los atributos a evaluar del pensamiento crítico, se propusieron 24 resultados de aprendizaje (relacionados con el contenido del examen de opción múltiple) de esta nueva herramienta y posteriormente los siete reactivos con 24 respuestas totales con base en el método propuesto por Olle Ten Cate - exámenes de razonamiento clínico basados en casos, CBCRT por sus siglas en inglés - (ten Cate, 2018) en el que se presenta un caso clínico que únicamente incluye sexo, edad y motivo de consulta principal. Posteriormente se proporciona información adicional que sirve como base para hacer una pregunta y entre paréntesis se pide al estudiante que seleccione determinada cantidad de respuestas de un área específica incluida en una lista con aproximadamente 60 opciones divididas en 4 áreas diferentes - diagnóstico, diagnóstico diferencial, estudios de laboratorio y tratamiento-, en este caso, por el nivel académico de los estudiantes se adaptó y se modificaron las áreas por “efectos en el metabolismo y la digestión”, “información obtenida en la historia clínica”, “vías metabólicas”, “hormonas y neurotransmisores”.

La revisión de la herramienta se dividió en tres fases. En la primera se convocó a seis miembros del CV2 y con base en sus observaciones se hicieron modificaciones y se envió a tres miembros más. Posteriormente se envió a los médicos residentes para que respondieran la herramienta con la intención de que su acercamiento fuera como el de los estudiantes y una vez que respondieron se les envió el material completo para que hicieran sus observaciones.

Una vez aprobadas ambas herramientas por los comités correspondientes, se estableció el punto de corte con el método de Angoff (Assessment Strategies Inc, 2014) y se invitó a 41 estudiantes (muestreo no probabilístico por conveniencia) de primer año que hubieran cursado los primeros tres bloques de la asignatura a que respondieran la prueba piloto para observar el comportamiento de la herramienta diseñada.

Para la prueba piloto se utilizó zoom y google forms. Durante la aplicación estuvieron presentes dos profesores de la asignatura y 30 estudiantes a quienes se les pidió que mantuvieran su cámara encendida todo el tiempo y se les envió por medio del chat la liga para la apertura del examen. Para asegurar la comprensión de las instrucciones y las evidencias de validez del proceso de respuesta para la herramienta de pensamiento crítico, resolvimos un caso clínico de prueba proyectando la pantalla.

Veintinueve estudiantes respondieron los dos exámenes - uno de ellos tuvo problemas de conexión y no pudo enviar las respuestas del examen de pensamiento crítico - y con base en ello se realizó la psicometría de acuerdo con la Teoría Clásica del Test (Chávez Álvarez y Saade Hazin, 2010) utilizando Excel para obtener por cada reactivo: grado de dificultad (proporción de sustentantes que aciertan al reactivo), índice de discriminación (diferencia entre proporción de respuestas correctas del grupo superior y el grupo inferior), punto biserial (correlación entre el reactivo y la calificación total del estudiante), varianza y α de Cronbach (estimación de la fiabilidad de la prueba).

Consideraciones éticas

Este proyecto cuenta con la aprobación por parte del Comité de Ética: FM/DI/118/2018, y fue aprobado el 5 de febrero de 2019.

La participación fue voluntaria, manteniendo en el anonimato la identidad de los participantes. Se les hizo entrega de una carta de invitación, un consentimiento informado, un acuerdo de confidencialidad y un aviso de privacidad. Asimismo, se les informó que eran libres de retirarse en el momento que consideraran necesario, aun si su participación no había concluido y que el resultado no afectaría las calificaciones de los estudiantes.

Resultados

Inicialmente consideramos que la aplicación de ambos exámenes duraría aproximadamente tres horas; sin embargo, el examen de conocimientos generales duró una hora y el de pensamiento crítico, 35 minutos. En total la aplicación duró una hora con 45 minutos tomando en cuenta el tiempo de las instrucciones. Al final de la aplicación se hicieron preguntas para que los estudiantes comentaran su experiencia: ¿Tuvieron alguna complicación? ¿Cuál es su opinión sobre este tipo de examen? ¿Respondieron todas las preguntas respetando el número de respuestas que se les pedían? ¿Les parece que es pesado responder los dos exámenes seguidos?

Algunas de las respuestas obtenidas fueron:

‘Creo que está bien porque [...] no nos están preguntando enzimas o cosas de relleno sino que es más de aplicación y de ver realmente para que se usa en cada ruta [...] son preguntas que siguen siendo de opción pero nos retan un poquito más a ver como van funcionando las vías metabólicas y las hormonas y todo lo que nos preguntaron. Entonces me parece que es un examen más completo’

‘[...] en algunas preguntas quería poner otras opciones más pero sólo me tenía que limitar a las que nos pedías’

‘[...] como que en la tercera parte dije: ¿cuántos años tenía y qué tenía? [...] me tuve que regresar a leer cuál era el planteamiento inicial’

‘[...] está bastante bien porque [...] sentí que no nos preguntan así nada más como para vomitar información sino que realmente nos están haciendo aplicar lo que hemos aprendido hasta ahora. Es como un poquito más complejo que incluso nos va a servir más a futuro porque realmente cuando estés frente a un paciente no te va a preguntar qué enzima le está fallando [...]’

En la psicometría obtuvimos un α de Cronbach de 0.83 para el examen de conocimientos generales y de 0.46 para el examen de pensamiento crítico (Tabla 1) e identificamos 12 reactivos (5 del examen de conocimientos generales y 7 del examen de pensamiento crítico) que se tienen que revisar (Tabla 2 y Tabla 3) ya que no cumplen con los criterios que establecimos para ser seleccionados: $GD > .20$, $ID > .15$, punto biserial $> .15$ (Hernández Arias, Martínez Lloreda y Hernández Lloreda, 2014).

Conclusiones

En este momento no podemos hacer inferencias sobre los resultados obtenidos ya que se trata de una prueba piloto, sin embargo, es importante mencionar que se tuvo especial cuidado en el proceso para contar con evidencias de validez de contenido, proceso de respuesta, estructura interna y relación con otras variables. Con respecto al α de Cronbach es importante mencionar que al ser calculada con base en la varianza, tener una muestra pequeña de estudiantes y ser un examen que los estudiantes no conocían era esperado tener un valor tan bajo. Por lo anterior los datos estadísticos no son concluyentes y estamos en espera de la aplicación en una población más grande para observar mejor el comportamiento.

Los reactivos que identificamos fuera de los parámetros establecidos serán revisados por el CR, CD y nuevamente por los CV y se tomarán en cuenta los comentarios de los estudiantes para mejorar el instrumento.

Pese a las limitaciones de los resultados, hasta ahora podemos concluir que es una herramienta con alta aceptación por parte de los estudiantes, útil desde su punto de vista, favorece la integración y aplicación de los conocimientos conceptuales y no representa problemas de tiempo ya que los exámenes aplicados tradicionalmente en la facultad, tienen una duración aproximada de dos horas.

Tablas y figuras

Tabla 1. Comparativo de ambos exámenes

Ex. Conocimientos generales		Ex. Pensamiento crítico	
Promedio	64.9	Promedio	49.1
Mediana	61.7	Mediana	45.8
Calif. Más baja	38.2	Calif. Más baja	29.2
Calif. Más alta	92	Calif. Más alta	70.8
GD	64.9	GD	49.1
Alfa de Cronbach	.83	Alfa de Cronbach	.46*

Tabla 2. Psicometría del examen de conocimientos generales

Examen de conocimiento generales							
Reactivo	GD	ID	Punto biserial	Reactivo	GD	ID	Punto biserial
36							
1*	53.3	-0.22	-0.2	18	56.7	0.56	0.44
2	73.3	0.44	0.36	19	66.7	0.67	0.5
3	73.3	0.44	0.35	20	70.0	0.44	0.48
4	36.7	0.56	0.55	21	63.3	0.56	0.49
5	53.3	0.33	0.29	22	73.3	0.33	0.42
6	70.0	0.44	0.37	23*	90.0	-0.11	-0.15
7*	16.67	0.11	0.12	24	70.0	0.33	0.24
8	60.0	0.56	0.47	25	70.0	0.33	0.38
9	73.3	0.56	0.51	26	80.0	0.22	0.34
10	70.0	0.44	0.25	27*	23.3	0.11	0.09
11	86.7	0.33	0.35	28	50.00	0.33	0.28
12	66.7	0.78	0.68	29	60.00	0.56	0.44
13*	73.3	0.00	0.11	30	76.7	0.33	0.44
14	70	0.56	0.50	31	60.0	0.56	0.43
15	83.3	0.44	0.45	32*	93.33	0.11	0.16
16	83.3	0.33	0.31	33	73.3	0.22	0.23
17	56.67	0.22	0.26	34	30.0	0.33	0.31

GD: grado de dificultad. *Requieren revisión
ID: índice de discriminación. * Requieren revisión

Tabla 3. Psicometría del examen de pensamiento crítico

Examen de conocimiento generales							
Reactivo	GD	ID	Punto biserial	Reactivo	GD	ID	Punto biserial
1*	82.8	0.13	0.28	13	51.72	0.31	0.26
2*	31.03	0.08	-0.03	14	24.14	0.33	0.21
3	68.9	0.39	0.36	15	31.03	0.33	0.12
4	75.9	0.63	0.57	16*	6.9	0.11	0.21
5	51.72	0.42	0.24	17	31.03	0.08	0.07
6	55.17	0.67	0.29	18	68.9	0.39	0.39
7	27.59	0.44	0.41	19	89.7	0.25	0.21
8	48.28	0.40	0.31	20	34.48	0.54	0.55
9*	68.97	0.14	0.23	21	44.83	0.53	0.47
10*	72.41	0.03	0.11	22	86.21	0.38	0.42
11	58.62	0.51	0.40	23	24.34	0.08	0.15
12*	27.59	0.08	0.12	24	17.24	0.21	0.28

GD: grado de dificultad. *Requieren revisión
ID: índice de discriminación. *Requieren revisión

Referencias

- Chávez Álvarez, C., & Saade Hazin, A. (2010). Fundamentos técnicos del análisis de reactivos. En *Procedimientos básicos para el análisis de reactivos* (págs. 49-97). Ciudad de México: Centro Nacional de Evaluación para la Educación superior.
- Lane, S., Raymond, M., Haladyna, T., & Downing, S. (2016). Test development process. En *Handbook of test development* (págs. 3-18). Nueva York: Routledge.

- Carrasco Cursach, J. (2018). Medición del desarrollo del pensamiento crítico en estudiantes chilenos/as de educación superior (Tesis de Maestría). Recuperada de <http://repositorio.udec.cl/handle/11594/3062>. 1-92.
- Castañeda Rivas, M. (2015). La enseñanza-aprendizaje por competencia. *Revista de la Facultad de Derecho*, 243-265.
- Assessment strategies Inc. (2014). *The Angoff method of standard setting*. Canadá.
- Assessment Strategies Inc. (2014). *The Angoff method of standard setting*. Canadá.
- Facione, P. A. (1998). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*.
- Facultad de Medicina. (2018). *Bioquímica y Biología Molecular: asignatura biomédica - primer año*. Ciudad de México.
- Fortoul Van der Goes, T., Morales López, S., Muñoz Comonfort, A., Jacobo Méndez, A., Varela Ruíz, M., & Rodríguez Lara, V. (2012). Retención de los conocimientos básicos en cinco generaciones de alumnos que terminaron los dos primeros años del plan único de la carrera de médico cirujano en la Facultad de Medicina, UNAM (2007-2011). *Revista de Investigación Médica*, 170-175.
- Hernández Arias, M., Martínez Lloreda, M., & Hernández Lloreda, M. (2014). La Teoría Clásica de los Tests II: puntuaciones, análisis de elementos, consideraciones finales. En *Psicometría* (págs. 69-100). Madrid: Alianza.
- Kuiper, R. (2004). Promoting cognitive and metacognitive reflective reasoning skills in nursing practice: self-regulated learning theory. *Issues and Innovation in nursing education*, 381-391.
- Makary, M., & Daniel, M. (2016). Medical error - the third leading cause of death in the US. *BMJ*, 1-5.
- Noll, E., Key, A., & Jensen, G. (2001). Clinical reasoning of an experienced physiotherapist: insight into clinician decision-making regarding low back pain. *Physiotherapy Reserach International*, 40-51.
- Paul, S. (2014). Assessment of critical thinking: A Delphi study. *Elsevier*, 1357-1360.
- Rivera Jiménez, J., Flores Hernández, F., Alpuche Hernández, A., & Martínez González, A. (2016). Evaluación de reactivos de opción múltiple en medicina. Evidencia de validez de un instrumento. *Revista de Investigación Médica*, 8-15.
- Sánchez Mendiola, M. (2018). La evaluación del aprendizaje de los estudiantes: ¿es realmente tan complicada? *Revista Digital Universitaria*, 1-18.
- Sánchez Mendiola, M., Durante Montiel, I., Morales López, S., Lozano, R., Martínez González, A., & Graue Wiechers, E. (2011). Plan de Estudios 2010 de la Facultad de Medicina de la Universidad Nacional Autónoma de México. *Gaceta Médica de México*, 152-158.
- Sadi, Ö. (2013). Case Study of How Turkish University Students Improve Their Biochemistry Achievement. *Higher Education Studies*, 52-67.
- ten Cate, O. (2018). Assessment of Clinical Reasoning Using the CBCR Test. En O. ten Cate, E. J.F.M., & S. Durning, *Principles and Practice of Case- based Clinical Reasoning Education: A method for preclinical students* (págs. 85-94). Holanda: Springer Open.
- Viesca Treviño, C., Ponce de León, M., & Sánchez Mendiola, M. (Sin fecha). El ejercicio actual de la Medicina. *Razonamiento clínico*, (pág. (Sin fecha)). Ciudad de México.
- Zayapragassarazan, Z., Menon, V., Kar, S., & Batmanabane, G. (2016). Understanding Critical Thinking to Create Better Doctors. *Journal of Advances in Medical Education and Research*, 9-13.