



Universidad Nacional Autónoma de México
Programa de Maestría y Doctorado en Ciencias Médicas, Odontológicas y de la
Salud

Campo del conocimiento:
Ciencias Sociomédicas (Educación en Ciencias de la Salud)

**Diseño de una herramienta para evaluar el pensamiento crítico en estudiantes de
Medicina**

Modalidad de graduación
Tesis

Que para optar por el grado de
Maestro en Ciencias

Presenta:
Oliva Briz Tena

Tutor:
Teresa I. Fortoul Van der Goes
Facultad de Medicina

Comité tutor:
Juan Pablo Pardo Vázquez
Facultad de Medicina
Armando Muñoz Comonfort
Facultad de Medicina

Ciudad Universitaria, CDMX, septiembre, 2021.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Dedicatoria

A mi esposo, sin ti no hubiera nacido la loca idea de emprender este viaje. Gracias por la paciencia y el apoyo, por acompañarme en los desvelos y hacer que fuera todo más sencillo.

A mi mamá y a mi papá que confiaron siempre en que lo lograría. Me dieron su apoyo incondicional y me motivaron para llegar hasta el final.

A mi hermana Jimena que prácticamente fue como mi tutora, gracias por resolver todas mis dudas y sacrificar tu sueño para ayudarme siempre.

A mi hermana Vanessa por motivarme siempre y hacer que este tiempo fuera divertido con sus ocurrencias.

Su apoyo y su amor me motivó a llegar hasta aquí. Sin ustedes nunca lo habría logrado. Han estado conmigo en todo momento y no fue fácil, pero lo logramos. Los amo, gracias por todo.

Agradecimientos

A la Mtra. Deya y la Dra. Tere quienes que no solo me acompañaron y guiaron académicamente sino que me brindaron su apoyo moral y estuvieron en los momentos más difíciles.

Al Dr. Juan Pablo y el Mtro. Armando que me guiaron, me acompañaron y enriquecieron con sus conocimientos.

A mis profesores que se han convertido en amigos, Héctor, Diego e Iván por su tiempo y disposición en todo el proyecto.

Al Dr. Nicolás y a Vero que además de ser excelentes profesores me ayudaron a superar uno de los mayores retos que he tenido en mi vida.

A mis amigos Moni, Alma, Fer y Kevin que creyeron en mí y me motivaron todos los días para continuar.

A mis maestros Melchor, Carlos, Adrián, Guadalupe, Maura, Bernardo, Isaías, Manuel, Nan, Paty, Alan y Nancy que me mostraron un mundo desconocido y sembraron en mí el deseo de conocer cada vez más sobre él.

A mis compañeros Mariana, Carlos, Angie, Carmen, Jesús, Kwei y Héctor por enseñarme tanto y hacer tan ameno este camino.

ÍNDICE

ÍNDICE	4
RESUMEN	7
CAPÍTULO 1. INTRODUCCIÓN	9
CAPÍTULO 2: PENSAMIENTO CRÍTICO	14
DEFINICIÓN DE PENSAMIENTO CRÍTICO	14
HABILIDADES DEL PENSAMIENTO CRÍTICO	15
IMPORTANCIA DEL PENSAMIENTO CRÍTICO	18
PENSAMIENTO CRÍTICO EN MEDICINA	19
PENSAMIENTO CRÍTICO Y CIENCIAS BÁSICAS	21
CAPÍTULO 3: EVALUACIÓN	24
DEFINICIÓN DE EVALUACIÓN	24
CARACTERÍSTICAS DE UN MÉTODO DE MEDICIÓN EFECTIVO	27
VENTAJAS Y DESVENTAJAS DE LOS EXÁMENES DE OPCIÓN MÚLTIPLE	28
EVALUACIÓN DE LA CALIDAD DE LOS REACTIVOS DE OPCIÓN MÚLTIPLE	33
EVALUACIÓN DEL PENSAMIENTO CRÍTICO	35
MÉTODOS UTILIZADOS PARA LA EVALUACIÓN DEL PENSAMIENTO CRÍTICO	37
MÉTODO PARA EVALUAR EL PENSAMIENTO CRÍTICO CON BASE EN CASOS CLÍNICOS	41

CAPÍTULO 4. PLANTEAMIENTO DEL PROBLEMA, JUSTIFICACIÓN Y OBJETIVOS DE LA INVESTIGACIÓN	46
PLANTEAMIENTO DEL PROBLEMA	46
JUSTIFICACIÓN	47
PREGUNTA DE INVESTIGACIÓN	48
OBJETIVO GENERAL	48
OBJETIVOS ESPECÍFICOS	48
CAPÍTULO 5: MÉTODO	49
DISEÑO DE INVESTIGACIÓN	49
PROCEDIMIENTO	49
ANÁLISIS ESTADÍSTICO	59
CAPÍTULO 6: RETOS ANTE LA PANDEMIA POR SARS-COV-2	60
CAPÍTULO 7: RESULTADOS	62
CAPÍTULO 7: DISCUSIÓN	72
CAPÍTULO 8: CONCLUSIONES	76
LIMITACIONES Y PERSPECTIVAS	77
REFERENCIAS	78
ANEXO 1 RECURSOS Y CONSIDERACIONES ÉTICAS	90

ANEXO 2 PLAN GENERAL	96
EXAMEN DE CONOCIMIENTOS GENERALES	96
EXAMEN DEL PENSAMIENTO CRÍTICO	102
ANEXO 3 CALIDAD DE LOS REACTIVOS	108
ANEXO 4 PSICOMETRÍA	116
EXAMEN DE CONOCIMIENTOS GENERALES	116
EXAMEN DE PENSAMIENTO CRÍTICO	117

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 error médico que en algunos países supone la tercera causa de muerte, por lo que es necesario tener evidencia de 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 con base en 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, familiares, urgenciólogos, internistas, nefrólogos y estudiantes del primer año de la licenciatura de Médico Cirujano.

Resultados: se realizó una prueba piloto con 29 estudiantes. Se obtuvo un alfa de Cronbach de 0.46 que se atribuyó al tamaño de la muestra y la poca familiaridad con el instrumento. La herramienta tuvo buena aceptación por parte de los estudiantes,

quienes consideraron que favorece la integración y aplicación de los conocimientos conceptuales.

CAPÍTULO 1. INTRODUCCIÓN

La educación es uno de los elementos fundamentales para el desarrollo de cualquier país. Su propósito es mejorar las condiciones de la sociedad en todas sus áreas mediante la **preparación** de las futuras generaciones, que estén capacitadas para resolver los problemas que se les presenten. Sin embargo, en la actualidad, para lograr lo anterior con mayor eficacia, es necesario que se abandone el sistema educativo tradicional en el que el estudiante se enfoca en la adquisición de conocimientos, muchas veces sin saber la utilidad de éstos; es decir, el aprendizaje de conceptos de forma memorística, que no necesariamente le ayudan a la solución de problemas (Castañeda, 2015; Riveros, 2015; Facione, 2016).

Si bien el sistema tradicional ha tenido cierta utilidad, el exceso de información a la que tenemos acceso hoy en día, la transición demográfica y los nuevos retos a los que nos enfrentamos nos exigen un cambio de modelo educativo. Como respuesta a estas nuevas necesidades, a finales del siglo XX, surge el proyecto Tuning, 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 (González, Wagenaar, & Benetione, 2004).

La educación guiada por competencias se sustenta en el constructivismo. Esta corriente de pensamiento coloca al estudiante como protagonista en el proceso de aprendizaje y este aprendizaje sería el producto de las interacciones sociales en las que se ve involucrado a lo largo de su vida (Anthony, 1996; Schunk, 2012b). En oposición, el modelo tradicional se basa en la teoría conductista, la cual ubica al docente como el principal actor en el proceso de enseñanza (Rojas & Díaz, 2012; Schunk, 2012a).

El enfoque por competencias ha tomado importancia en los distintos niveles de la educación. Un claro ejemplo es la Facultad de Medicina de la UNAM, que en el año 2010 propuso modificaciones al plan existente. El nuevo plan de estudios de la licenciatura de Médico Cirujano implementó un currículo mixto, estructurado por asignaturas y guiado por competencias, que plantea un sistema educativo que garantice la calidad de la educación y la formación de médicos altamente calificados al aplicar sistemas de evaluación válidos y confiables (Consejo Técnico de la Facultad de Medicina, 2010).

De acuerdo con Hunter y Epstein, la competencia se define como “el uso juicioso y habitual de la comunicación, conocimientos, habilidades técnicas, razonamiento clínico, valores y reflexión en la práctica diaria en beneficio de los individuos y comunidades que se atienden” (Epstein, 2007). En el ámbito de la medicina, la Asociación de Escuelas y Facultades de Medicina (AMFEM), siguiendo a Rychen y Hersh, define a las competencias como “la capacidad para responder con idoneidad a demandas complejas en entornos dinámicos que presentan incertidumbre, mediante desempeños sustentados en el mejor conocimiento disponible, con profesionalismo y ética” (Castañeda, 2015; González, Wagenaar, & Benetione, 2004).

Con base en lo anterior y con el objetivo de egresar médicos capaces de solucionar situaciones del quehacer médico actual, el H. Consejo Técnico de la Facultad de Medicina de la UNAM, a través de una propuesta de cambio al plan de estudios, incluyó las siguientes competencias para el plan de estudios 2010 (Consejo Técnico de la Facultad de Medicina, 2010):

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.

Es evidente que las competencias que contiene el plan curricular no se basan únicamente en la adquisición de conocimiento teórico, sino que buscan desarrollar áreas psicosociales e inter e intrapersonales que mejoren el desempeño de los futuros médicos. Por ello, es importante la implementación de estas competencias para que cuando los estudiantes se enfrenten con situaciones desconocidas, propongan soluciones adecuadas.

El uso de las competencias no sólo tendrá un beneficio en el proceso de aprendizaje del estudiantado, sino que podría ser de ayuda para resolver un problema en el campo de la medicina. Makary y Daniel reportan que la tercera causa de muerte en Estados Unidos es el error médico (Makary & Daniel, 2016) y una de las razones podría ser la falta de pensamiento crítico, por lo que la deficiencia 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, & Batmanabane, 2016).

A partir de lo anterior, Harasym et. al. proponen que los estudiantes con pensamiento crítico tienen mejores habilidades para la solución de problemas y toma de decisiones, lo que hace que sean médicos más competentes (Harasym, Tsai, & Hemmati, 2008).

Aunado a la necesidad de fortalecer el conocimiento clínico y el pensamiento crítico en los estudiantes, es necesaria la aplicación de métodos de evaluación que ayuden a definir la adquisición de este elemento. Así, siguiendo la cita atribuida a Peter Drucker, “lo que se mide se puede mejorar” (Kaufman & Reichert, 2018) en este trabajo nos enfocaremos en la propuesta de una herramienta que sirva para medir el pensamiento crítico en los alumnos. Asimismo, nos ceñiremos a la visión de que la evaluación es uno de los rasgos más importantes de la educación, ya que ésta es fundamental para valorar el logro de los objetivos de aprendizaje (Jiménez, Flores, Alpuche, & Martínez, 2016) y con ello llevar a cabo un proceso de enseñanza y aprendizaje integral (Morales, Varela, & Núñez, 2015).

CAPÍTULO 2: EL PENSAMIENTO CRÍTICO

“Solo el pensamiento crítico conduce al desarrollo de la sociedad”

Ibiza Melián

A pesar de que muchos autores han propuesto una definición del pensamiento crítico y han analizado sus características, aún hay controversia con relación a su definición.

En 1987, la Asociación Filosófica Americana (APA por sus siglas en inglés, *American Philosophical Association*) convocó a un grupo de expertos para hacer una investigación sistemática sobre el estado del pensamiento crítico y la evaluación de éste. Con base en el método Delphi, se llegó a la conclusión de que el pensamiento crítico es 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). Cada una de estas propiedades del pensamiento crítico se encuentra explicada con mayor detalle en la Cuadro 1.

Cuadro 1. Habilidades del pensamiento crítico

Habilidades	Descripción	Actividades
Interpretación	Comprender y expresar el significado de experiencias, situaciones, eventos, juicios, creencias, reglas, procedimientos, etc.	Categorización. Decodificación del significado. Clarificación de conceptos.
Análisis	Identificar relaciones entre declaraciones, preguntas, conceptos, expresiones, etc.	Examinación de ideas. Detección y análisis de argumentos.
Evaluación	Evaluar la credibilidad de los argumentos que son resultado de la percepción o experiencia de una persona. Evaluar las relaciones de inferencia entre declaraciones, preguntas, conceptos, expresiones, etc.	Evaluar la credibilidad de demandas. Evaluar la calidad de argumentos que se utilizan e inducen o deducen razonamientos.
Inferencia	Identificar y asegurar los elementos que se requieren para crear una conclusión razonable. Formar conjeturas e hipótesis. Considerar información relevante y deducir las consecuencias provenientes de datos, evidencia, creencias, juicios, opiniones, conceptos, descripciones, etc.	Consulta de pruebas. Conjetura de alternativas. Obtención de conclusiones.
Explicación	Representación coherente de los resultados. Presentar argumentos convincentes y justificar el razonamiento mediante el cual se llegó a determinados resultados.	Descripción de métodos y resultados. Justificación de procedimientos, objetivos y explicaciones conceptuales, argumentaciones, etc.
Autorregulación	Actividad cognitiva de monitoreo autoconsciente. Se emplean habilidades de análisis y evaluación.	Autoexaminación. Autocorrección.

(Núñez, Ávila, & Olivares, 2017; Facione, 1998)

Halpern propone cuatro pilares para el desarrollo del pensamiento crítico (Halpern, 1998):

- Disposición del estudiante para esforzarse en pensar y aprender: implica hacer un esfuerzo mental consciente, tener compromiso, ser flexible, tener voluntad para la autocorrección y consciencia de la realidad social que debe mejorar.
- Enfoque de habilidades para el pensamiento crítico para que el estudiante comprenda cómo se determina una causa, reconozca y critique supuestos, analice la relación entre medios y fines, construya argumentos sólidos para fundamentar inferencias, evalúe los grados de probabilidad e incertidumbre, incorpore datos aislados a un marco mayor y use analogías para resolver problemas.
- Diseño de la capacitación por parte de los profesores: con la intención de que el estudiante no sólo reconozca las habilidades requeridas, sino que comprenda su utilidad y sea capaz de aplicarlas en situaciones nuevas.
- Metacognición: autoconciencia y planificación que guíe el uso de las habilidades del pensamiento crítico.

Aunado a lo anterior, se propone que los recursos intelectuales requeridos para el pensamiento crítico son (Bailin, Case, Coombs, & Daniels, 1999)

- Conocimiento previo: conceptos, valores y creencias. La profundidad de este conocimiento determina el grado en el que el estudiante será capaz de pensar críticamente.
- Conocimiento operativo de los estándares de buen pensar: práctica deliberada, argumentación y desarrollo de planes de acción.

- Conocimiento de conceptos críticos: aquellos que permiten al buen pensador hacer distinciones e interpretaciones.
- Heurísticas: estrategias o atajos de pensamiento que permiten llevar a cabo tareas específicas.
- Hábitos mentales: respeto por la verdad, actitud de investigación, imparcialidad, independencia, respeto a la investigación de otros, respeto por la autoría intelectual y ética.

Además de las habilidades específicas requeridas para el desarrollo del pensamiento crítico, como mencionan Paul y Elder (Paul & Elder, 2003), es necesario que un pensador crítico posea características como:

- Humildad intelectual: estar conscientes de sus limitaciones y no pretender saber más de lo que sabe.
- Empatía intelectual: ser capaz de ponerse en el lugar de otro para entenderlo.
- Autonomía intelectual: comprometerse a analizar su pensamiento basado en la evidencia.
- Integridad intelectual: ser honesto al evaluar sus ideas y los puntos de vista de los demás.
- Perseverancia intelectual: buscar la verdad, aunque haya dificultades y frustraciones.

- Confianza en la razón: confiar en que todos tienen la capacidad para aprender a pensar por sí mismos.
- Imparcialidad: abordar todos los puntos de vista a pesar de tener una opinión bien establecida.

Asimismo, este autor sugiere que para lograr el pensamiento crítico hay que pasar por 6 niveles y distingue 6 tipos de pensadores:

- Pensador irreflexivo: no está consciente de su propio pensamiento.
- Pensador retado: se enfrenta con problemas en su propio pensamiento.
- Pensador principiante: trata de mejorar, pero sin práctica regular.
- Pensador practicante: reconoce la necesidad de práctica regular.
- Pensador avanzado: avanza según su práctica.
- Pensador maestro: tiene buenos hábitos de pensamiento y estos se vuelven parte de su naturaleza.

Importancia del pensamiento crítico

Por tanto, se entiende que la meta ya no es la acumulación de información por parte de los estudiantes, característica del método tradicional de educación, sino el fortalecimiento de las habilidades del pensamiento crítico para que los alumnos sean capaces de aplicarlo en un contexto adecuado (Mafinejad, et. al., 2017). De igual manera, se considera de suma importancia la adquisición del pensamiento crítico

como eje central de una educación integral y de calidad. Sólo si un estudiante cuenta con la capacidad de reflexión y resolución de problemas, estará preparado para llevar a cabo su profesión de manera responsable (Bezanilla, Poblete, Fernández, Arranz, & Campo, 2018).

Siguiendo a Facione, el pensamiento crítico promueve la autonomía racional, la libertad intelectual y la investigación objetiva y basada en evidencias para la solución de una amplia gama de problemas, tanto personales como sociales (Facione, 1998). De acuerdo con el Plan de estudios 2010, los profesionales en la actualidad se enfrentan con situaciones no rutinarias y con problemas inéditos que demandan soluciones innovadoras y requieren capacidades intelectuales de orden superior (Consejo Técnico de la Facultad de Medicina, 2010), por lo que el desarrollo del pensamiento crítico es indispensable.

Pensamiento crítico en Medicina

Algunos autores consideran que el pensamiento crítico es una de las claves en el proceso del razonamiento clínico (Kuiper & Pesut, 2004) y que este último es la “facultad que permite resolver problemas médicos, extraer conclusiones y aprender de manera consciente de los hechos, estableciendo conexiones causales y lógicas necesarias entre ellos” (Villaruel & Bernal, 2014), mientras que otros refieren que el

pensamiento crítico aplicado a una situación clínica es el razonamiento clínico en sí (Noll, Key, & Jensen, 2001; Victor-Chmil, 2013) y es la base para la toma de decisiones que afectan la salud y la calidad de vida de los pacientes (Viesca, Ponce de León, & Sánchez). 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 utilizemos un método de evaluación basado en casos clínicos, el cual será mencionado más adelante.

Algunos estudios describen el procedimiento para lograr el razonamiento clínico y proponen que el proceso del pensamiento es una función integradora de dos sistemas, el intuitivo y el reflexivo. El primero está relacionado principalmente con heurísticas —ideas de asociación fácil y memorias vívidas para llegar a un juicio rápidamente— y es útil en situaciones familiares y cuando se necesita una acción inmediata (Sánchez & Martínez, 2014; Facione, 2016). Este tipo de razonamiento es automático y no requiere plena consciencia, por lo que normalmente es utilizado por especialistas (Viesca, Ponce de León, & Sánchez) con amplia experiencia en el reconocimiento de patrones.

El segundo, el reflexivo, se utiliza para hacer juicios cuando la situación es desconocida y se tiene más tiempo para resolver el problema. Asimismo, permite procesar conceptos abstractos para planear con anticipación y considerar opciones

cuidadosamente (Facione, 2016). Este modelo utiliza el pensamiento analítico; se caracteriza por ser lento, pero más consciente e implica la observación cuidadosa, obtención de información, exploración física, generación de hipótesis, correlación de los datos obtenidos y confirmación o rechazo de las hipótesis planteadas (Sánchez & Martínez, 2014; Villarroel & Bernal, 2014).

En la Universidad de Michigan y la Universidad de McMaster se realizó un estudio sobre la diferencia en los procesos de pensamiento entre un grupo de expertos y un grupo de estudiantes. Los investigadores demostraron que los expertos recuerdan información específica para la elaboración de diagnósticos; pese a ello, retienen menos información que los novatos. Asimismo, los expertos utilizan menos información de las ciencias básicas que los novatos y esto se atribuye a la separación evidente de las ciencias básicas y clínicas en los planes de estudio (Viesca, Ponce de León, & Sánchez; Kunina, Hautz, Knigge, Spies, & Ahlers, 2015) .

Pensamiento crítico y ciencias básicas

De acuerdo con varios autores, la comprensión de las ciencias básicas es un determinante importante en el éxito del diagnóstico a largo plazo (Woods, et. al., 2006; Boshuizen & Schmidt, 1992). Es decir, un estudiante que no tiene los conocimientos suficientes en la asignatura no será capaz de desarrollar las

habilidades del pensamiento crítico en esta área. Para que el estudiante aprenda, es necesario el conocimiento de conceptos, su internalización y posterior aplicación, así como una constante evaluación de su trabajo (Núñez, Ávila, & Olivares, 2017).

El pensamiento crítico en contextos particulares, como sería el de la asignatura de Bioquímica y Biología Molecular, demanda conocimiento específico (Facione, 1998) y tiene un rol 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 se puede cumplir si el estudiante está en contacto con casos clínicos durante su formación (Viesca, Ponce de León, & Sánchez).

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 y por lo tanto, también su pensamiento crítico lo es (Fortoul, et. al., 2012; Zayapragassarazan, Menon, Kar, & Batmanabane, 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).

Partiendo de lo anterior, es necesario favorecer el desarrollo del pensamiento crítico desde los primeros años de formación, introduciendo casos clínicos en la educación del alumno (Viesca, Ponce de León, & Sánchez). Para lograr el desarrollo de esta competencia, es necesario que los estudiantes sean capaces de identificar, seleccionar, recuperar y discernir información de diversas fuentes y así estar mejor preparados para desenvolverse en escenarios clínicos reales (Consejo Técnico de la Facultad de Medicina, 2010; McDonald, Howard, Schlump, & Plack, 2014).

La práctica con casos clínicos debe simular la aplicación del conocimiento en un contexto clínico y se recomienda que la presentación de los casos clínicos no esté relacionada con el tema que se está revisando, ya que esto sugiere un diagnóstico y puede restarle efectividad a esta estrategia. De igual forma, es recomendable vincular al estudiante con la clínica en etapas tempranas y con las ciencias básicas en etapas avanzadas de su formación, lo que les ayudará a generar hipótesis con fundamentos tanto bioquímicos como fisiológicos, entre otros (Viesca, Ponce de León, & Sánchez).

CAPÍTULO 3. EVALUACIÓN

De acuerdo con los estándares definidos por la *American Educational Research Association*, la *American Psychological Association* y el *National Council on Measurement in Education*, evaluación se define como “cualquier método sistemático para la obtención de información a partir de exámenes y otros recursos con el propósito de hacer inferencias relacionadas con las características de una persona, objeto o programa” (Downing & Haladyna, 2009b). Para Epstein, este elemento tiene tres objetivos: “optimizar las capacidades de los estudiantes, proteger a los pacientes identificando médicos incompetentes y proveer las bases para elegir a aquellos que solicitan entrenamiento avanzado” (Ferris & Flynn, 2015).

En el caso de la Facultad de Medicina, debido a la gran cantidad de alumnos que ingresan, 1561 para el periodo 2019-2020 (Fajardo, 2019), y a su aumento año con año, es complejo y difícil implementar una herramienta cuya aplicación requiera tiempo para evaluar estudiante por estudiante. De este modo, resulta más práctico crear herramientas que se puedan aplicar a grupos de estudiantes de manera simultánea, a través de los profesores de asignatura, y de esta forma dar un paso más hacia el cumplimiento de los objetivos del plan de estudios 2010. La implementación de herramientas de medición de opción múltiple resulta

conveniente y podría arrojar resultados satisfactorios si se elaboran evaluaciones bien estructuradas.

Si bien en esta investigación se aboga por los exámenes de opción múltiple como una buena forma de lograr nuestros objetivos, no debe dejarse de lado que todos los métodos de evaluación tienen fortalezas y debilidades y por ello existe una amplia gama de instrumentos que permiten evaluar las distintas competencias (Epstein, 2007; Downing & Haladyna, 2009b; Durante, Lozano, Martínez, Morales, & Sánchez, 2012; Ferris & Flynn, 2015). Así, el instrumento más adecuado de evaluación depende de aquello que se pretende evaluar (Jiménez, Flores, Alpuche, & Martínez, 2016). Para lograr este objetivo, Van der Vleuten propone cinco criterios que determinan la utilidad de cada método de evaluación: confiabilidad, validez, el impacto en la práctica y el aprendizaje futuro, aceptabilidad para los estudiantes y profesores y los costos (Epstein, 2007; Downing & Haladyna, 2009b).

Aunque todos los criterios anteriores son importantes, la validez es el *sine qua non* de todos los datos de evaluación y se refiere a las “pruebas presentadas para respaldar o refutar la interpretación asignada a los resultados de la evaluación”. En otras palabras, se busca que un instrumento mida lo que en realidad debe medir (Downing & Haladyna, 2009b; Jiménez, Flores, Alpuche, & Martínez, 2016; Sánchez, 2018; Areekkuzhiyl, 2019). Los estándares describen cinco tipos de validez: de contenido, del proceso de respuesta, de la estructura interna, de relación con otras

variables y de las consecuencias (Downing & Haladyna, 2009b; Jiménez, Flores, Alpuche, & Martínez, 2016; Sánchez, 2018).

Cabe destacar que existen tres dominios que deben ser evaluados en el estudiantado: el cognitivo, el afectivo y el psicomotor. El primero de ellos, y en el cual se basa este trabajo, incluye el conocimiento, la comprensión, la aplicación, el análisis, la síntesis y la evaluación (Khan & Aljarallah, 2011). Los otros dos dominios tienen que ver con actitudes, valores y apreciaciones y con habilidades de movimiento, respectivamente (Bloom, Engelhart, Frust, Hill, & Krathwohl, 1956).

Los exámenes de conocimiento cognitivo, es decir, aquellos que evalúan el aprendizaje asociado a habilidades mentales, son los más comunes e importantes en la educación médica. La modalidad escrita es la más utilizada y está dividida en dos grandes grupos: los de selección de respuesta y los de construcción de respuestas (Downing S., 2002).

Dentro de los exámenes de selección de respuesta se encuentran los de opción múltiple, los de relación de columnas (corta o extendida), los de verdadero o falso (con su variante múltiple) y los de opción alternativa (Downing & Yudkowski, 2009).

De acuerdo con Downing, cualquiera que sea el formato elegido —opción múltiple, relación de columnas, verdadero o falso u opción alternativa— las evaluaciones deben contar con ciertas características deseables (Downing S., 2002):

- Objetividad: los expertos en el contenido del examen deben estar de acuerdo con que la respuesta marcada como correcta sea en realidad la correcta.
- Propiedades de medida: su calificación debe ser emitida de acuerdo con escalas de puntuación que cuenten con evidencias de validez, sean reproducibles y correspondan con el nivel de conocimiento esperado por los estudiantes.
- Evidencias de validez: la de contenido es la más importante y es por ello que debe contarse con evidencias sobre la forma en la que se seleccionó el contenido y la experiencia de los elaboradores de reactivos.
- Reproducibilidad: teóricamente, una segunda administración del mismo examen a los mismos estudiantes debería emitir los mismos puntajes. Dos métodos utilizados para estimar esta característica son el coeficiente de Kuder-Richardson y el coeficiente de alfa de Cronbach. La Teoría de generalizabilidad permite analizar diversas fuentes de error y estima cada una por separado (Zúñiga & Montero, 2007; DOF, 2017; Martínez, Hernández, & Hernández, 2014b).
- Correspondencia de las calificaciones con el logro de aprendizaje: altos puntajes en el examen deben indicar alto dominio del tema, mientras que bajos puntajes deben indicar poco dominio del tema.

- Evidencias para sustentar las inferencias: debe haber suficientes evidencias de que las inferencias hechas sobre los puntajes son adecuadas, lo cual se logra con las evidencias de validez.
- Diligencia: evidencia de que hubo suficiente cuidado para asegurar la legitimidad y exactitud de los puntajes y las inferencias hechas sobre estos.
- Puntajes de pase o no pase: debe haber un proceso adecuado para establecer el punto de corte para determinar cuál es la calificación aprobatoria (DOF, 2017).
- Eficiencia: en tiempo y costo, así como facilidad para elaborarse, administrarse y puntuarse.

Cualquier evaluación que cuente con los elementos anteriores podrá considerarse un método de medición efectivo y logrará dar un resultado satisfactorio para conocer el rendimiento académico y el conocimiento teórico de los alumnos.

Ventajas y desventajas de los exámenes de opción múltiple

Los exámenes de selección de respuesta, como cualquier método de evaluación, cuentan con fortalezas y debilidades. Dentro de las ventajas que tiene este tipo de evaluación, podríamos enlistar las siguientes: el nivel de concordancia que existe entre los expertos sobre la respuesta correcta es alto, son objetivos, todos los tipos de reproducibilidad se obtienen fácilmente, tienen evidencias de validez cuando se elaboran correctamente y permiten hacer pruebas piloto. No obstante, hay quienes

han hecho fuertes críticas a su efectividad por tener limitaciones como la probabilidad de respuesta al azar, la aparición de señales sobre la respuesta correcta dentro de la formulación de la pregunta y la presencia de contenido trivial que únicamente evalúe la capacidad de recordar (Downing, 2002; Downing & Haladyna, 2009b).

Para disminuir las desventajas y contar con las características deseables de los exámenes de opción múltiple, Downing propone doce pasos que maximizan las evidencias de validez de las inferencias emitidas con los resultados de este tipo de exámenes (Downing S., 2006):

- Establecer un plan general en donde se establezcan diversos factores como la definición del constructo a evaluar, el propósito de la prueba, la interpretación que se hará con los resultados, el formato (en este caso selección de respuesta - opción múltiple), la modalidad (papel y lápiz o digital), una línea de tiempo en donde se establezcan fechas estimadas para cada paso, responsables de cada tarea, plan para el control de la seguridad del examen como un acuerdo de confidencialidad, el método que se utilizará para establecer el punto de corte y el modelo propuesto para la psicometría del examen (Teoría clásica del Test, Teoría de Generalizabilidad, Teoría de Respuesta al Ítem) .

- Definir el contenido del examen, el cual tendrá variaciones dependiendo del propósito de la prueba —evaluación diagnóstica, formativa o sumativa—, las consecuencias —altas, bajas o medias— de las decisiones tomadas con base en los resultados y el grado en el que pueden defenderse las evidencias requerida para la toma de decisiones. El contenido puede ser definido con base en programas operativos, documentos curriculares, libros de texto, entre otros.
- Elaborar una tabla de especificaciones que contenga como mínimo el formato y tipo de examen que se utilizará —selección de respuesta - opción múltiple—, el número total de ítems a elaborar, la taxonomía que se utilizará (Ej. Taxonomía de Bloom), si el examen incluirá contenido visual o auditivo, la puntuación —1: correcta, 0: incorrecta—, la forma en que se interpretarán los resultados y el tiempo límite para responder cada reactivo.
- Elaborar reactivos con base en la tabla de especificaciones. Los elaboradores deberán estar entrenados y no es suficiente que sean expertos en la materia, sino que deberán contar con habilidades en la elaboración de reactivos. En caso de que los elaboradores no tengan conocimiento al respecto, es recomendable que los novatos trabajen en parejas con un guía que tenga experiencia, pues, además de cuidar el contenido, es importante tomar en cuenta la diversidad cultural, religiosa y étnica, así como procurar la equidad de género y evitar herir susceptibilidades. De la misma manera se recomienda que el contenido sea

congruente con los resultados de aprendizaje, se evite elaborar preguntas extensas o con contenido trivial, expresiones extremas y dobles negaciones.

- Ensamblar y dar formato. Resulta importante disminuir la carga cognitiva — ítems completos en una sola página—, cuidar que exista un balance entre las opciones correctas (A=B=C=D), ordenar las respuestas jerárquicamente, ya sea ascendente o descendentemente. Procurar la seguridad en este punto es importante ya que más personas tienen acceso a los ítems.
- Imprimir y empaquetar. El autor menciona la importancia de establecer políticas de seguridad ya que la información está disponible para múltiples personas que quizá no son parte del comité elaborador de reactivos que han acordado confidencialidad. De igual modo, se recomienda eliminar el exceso de material impreso, llevar a cabo una revisión final del examen y corregir en caso necesario, así como verificar que el material sea visible, claro y ordenado.
- Administrar el examen y planear con anticipación el tiempo que durará, así como cuidar las condiciones ambientales. De igual modo, en este punto se tendrían que tomar en cuenta otros requerimientos como explicar las instrucciones a los aplicables y establecer políticas de seguridad para el transporte, recepción, distribución, recolección y transportación de los exámenes después de que han sido aplicados. En este punto es necesario tomar en cuenta la capacitación de los aplicadores y nombrar a un encargado para la supervisión del proceso.

- Puntuar las respuestas. Es importante llevar a cabo un análisis adecuado por medio de parámetros psicométricos. Para ello existen distintos modelos que pueden ser útiles como la Teoría clásica del Test —la cual parte del supuesto de que el puntaje obtenido por un estudiante corresponde a la suma del puntaje verdadero y un error aleatorio, sin embargo no estudia las fuentes de error—, La Teoría de Generalizabilidad —que reconoce que existen distintas fuentes de error y estima cada una por separado— y la Teoría de Respuesta al Ítem — que modela la probabilidad de respuesta en función de las características del estudiante y del reactivo— (DOF, 2017).
- Establecer puntos de corte con métodos sistemáticos como el de Angoff o el método de Ebel. En el primero, los expertos determinan la probabilidad de que un estudiante mínimamente competente responda correctamente cada ítem, posteriormente se suman y se promedian estas probabilidades y se obtiene el puntaje de aprobación. Por otro lado, el segundo se distingue del anterior en que los ítems se categorizan de acuerdo con su relevancia —esencial, importante, aceptable y cuestionable— y dificultad —fácil, medio y difícil—, se asignan las probabilidades de responder correctamente y se ponderan para establecer el puntaje mínimo de aprobación (DOF, 2017).
- Reportar los resultados con puntualidad una vez que se haya verificado que no hay errores en el reporte. El idioma debe ser comprensible para los receptores de

los resultados y se debe evitar el mal uso de estos al hacer inferencias que no concuerden con los objetivos de la evaluación. Asimismo, es necesario especificar el punto de corte y la escala utilizada.

- Generar un banco de ítems con los reactivos previamente elaborados. Producir reactivos útiles es un proceso largo y costoso, por lo que después de llevar a cabo el análisis psicométrico, se recomienda que aquellos ítems de buena calidad sean resguardados y almacenados de manera segura para evitar perderlos. Se deberán clasificar y guardar con información sobre su contenido, el nivel cognitivo que evalúan, su formación histórica e índices de dificultad y discriminación.
- Elaborar un reporte técnico en donde deberían incluirse aspectos importantes sobre el desarrollo del examen, evidencias de validez y recomendaciones.

Evaluación de la calidad de los reactivos de opción múltiple

Además de lo anterior, Jiménez y colaboradores proponen un instrumento para evaluar la calidad en la elaboración de los reactivos de opción múltiple, de acuerdo con los Standards for Educational and Psychological Testing, que consta de cuatro factores compuestos por distintos ítems que deben ser contestados para lograr un método de evaluación confiable (Jiménez, Flores, Alpuche, & Martínez, 2016):

- Comprensión del reactivo:
 - ¿La cantidad de texto en el tallo es adecuada para su comprensión?

- ¿La pregunta o instrucción se encuentra redactada con claridad?
 - ¿El reactivo cuenta con una gramática, puntuación y ortografía correctas?
 - ¿El tallo del reactivo plantea la idea central?
- Contenido del reactivo:
- ¿El reactivo presenta un solo resultado de aprendizaje?
 - ¿El reactivo presenta un solo contenido temático?
 - ¿La semántica utilizada está de acuerdo con el contenido del programa académico?
- Precisión del reactivo:
- ¿El reactivo cuenta únicamente con una respuesta correcta?
 - ¿Las opciones son independientes entre sí?
 - ¿El contenido evaluado está en relación con la especificación del reactivo?
- Redacción de opciones de respuesta
- ¿Las opciones son similares en cuanto a estructura gramatical, contenido y extensión?
 - ¿Las opciones evitan dar pistas sobre la respuesta correcta?
 - ¿Los distractores son plausibles, es decir, no se descartan por inferencia lógica o sentido común?
 - ¿El reactivo cuenta con tres o cuatro opciones de respuesta?

Teniendo en cuenta la información anterior y las distintas características que se deben considerar para la elaboración de métodos de evaluación efectivos, es necesario trabajar para la propuesta de una herramienta de medición que cumpla con los objetivos que aquí se trabajaron.

Evaluación del pensamiento crítico

El principal objetivo de este trabajo es desarrollar una herramienta de evaluación que nos permita conocer el nivel del pensamiento crítico de nuestros alumnos para saber en qué áreas se debe mejorar. Si logramos evaluar este elemento y mejorar los puntos en los que se está fallando, apoyaremos con la formación de médicos capaces que en un futuro tratarán con vidas humanas y estarán calificados para llegar a diagnósticos más precisos (Khan & Aljarallah, 2011).

Para alcanzar lo anteriormente mencionado, es necesario definir claramente los aspectos del pensamiento crítico a evaluar (Facione, 1998). De acuerdo con el plan de estudios 2010, las áreas del pensamiento crítico que se espera que el estudiante de primer año logre, son (Consejo Técnico de la Facultad de Medicina, 2010):

- 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.

En este proyecto se plantea el desarrollo de un instrumento de medición que estará enfocado en evaluar únicamente los dos últimos puntos. Estos coinciden con los rubros de interpretación y análisis de la definición de Facione vistos anteriormente en el Cuadro 1, que empatan con las actividades propuestas por el mismo autor: clarificar, examinar ideas y detectar y analizar argumentos.

A continuación, se enlistan las definiciones, en aras de claridad, de las actividades contenidas en el proyecto Delphi que serán utilizadas en este proyecto (Facione, 1998):

- Categorización: reconocer un problema y definir su carácter sin prejuicios, así como determinar un camino para clasificar y subclasificar información,
- Clarificar: parafrasear o hacer explícito a través de la descripción o analogía los significados de palabras, conceptos, comportamientos, números, signos. Su fin es la descripción para remover cualquier confusión.

- Examinar ideas: reconocer problemas, determinar sus componentes e identificar sus relaciones conceptuales
- Detectar argumentos: determinar si las descripciones, preguntas o gráficos dados expresan o no razones que sustenten su punto de vista.
- Analizar argumentos: identificar y diferenciar las premisas que permitan sustentar una conclusión tras el análisis de una expresión, opinión o punto de vista dados.

Es importante mencionar que medir el pensamiento crítico es complejo, ya que deben tomarse en cuenta diversos procesos mentales y la mayoría de los instrumentos desarrollados corresponden a contextos distintos al ámbito latinoamericano (Calle, 2013). Aunado a lo anterior, no existe un consenso sobre la evaluación del pensamiento crítico (Mafinejad, y otros, 2017).

Métodos utilizados para la evaluación del pensamiento crítico

Como se ha comentado anteriormente, existe controversia con respecto a la definición del pensamiento crítico y como resultado de esto, es complicada su evaluación, mas no imposible. A continuación, se mencionan algunos ejemplos de herramientas con este objetivo.

El California Critical Thinking Skills Test (CCTST) se ha utilizado durante 30 años en distintos países ya que está disponible en 18 idiomas, incluido el español, para evaluar habilidades generales de razonamiento, análisis, interpretación, evaluación, inferencia, aritmética y otros aspectos del pensamiento crítico general en estudiantes con distintas edades, niveles educativos y campos de interés. El examen consta de 34 reactivos de opción múltiple que incluye escenarios cotidianos y varían en dificultad y complejidad (Insight Assessment, 2021b). Tiene una duración de aproximadamente 45 minutos, los resultados se emiten en cuanto termina el examen y tiene un costo entre 15 y 30 dólares (ETSU, 2020). Existe también una variante de este examen enfocada en Ciencias de la salud, el Health Science Reasoning Test (HSRT) dirigido a estudiantes de Medicina, Odontología y Enfermería entre otros; sin embargo, no es necesario tener conocimientos sobre Ciencias de la salud, la información correspondiente a cada escenario es provista en cada caso clínico (Insight Assessment, 2021b) y tiene un costo aproximado de 40 dólares por estudiante (RCSJ, 2021).

Cox y McLaughlin estudiaron la relación entre la calificación obtenida en el HSRT y el desempeño académico de los estudiantes de la Escuela de Farmacia encontrando resultados estadísticamente significativos únicamente en la sección de ética (Cox & McLaughlin, 2013).

El Halpern Critical Thinking Assessment (HCTA) consta de 25 escenarios de la vida diaria (educación, salud y trabajo entre otros) en los que inicialmente se responden reactivos de construcción de respuesta y posteriormente de opción múltiple para evaluar habilidades de razonamiento verbal, argumentación, pensamiento como prueba de hipótesis, probabilidad e incertidumbre y solución de problemas. Se ha utilizado en estudiantes de distintos niveles educativos y nacionalidades (Butler, et. al., 2012). La versión española (PENCRISAL) se ha aplicado en estudiantes de bachillerato y estudiantes universitarios encontrando un alfa de Cronbach desde .34 hasta .635 dependiendo de la dimensión evaluada (Nieto, Saiz, & Orgaz, 2009).

El examen de pensamiento crítico PENCRISAL consta de 35 reactivos de construcción de respuesta en español en los que se evalúan componentes del pensamiento crítico general como razonamiento deductivo, inductivo y práctico, toma de decisiones y solución de problemas. No existe un tiempo límite para resolverlo, sin embargo, el tiempo promedio que tardan los estudiantes es de 60 - 90 minutos y se ha aplicado en estudiantes de 18 a 53 años. Cada reactivo puede tener una calificación de cero puntos (si la respuesta es incorrecta), un punto (si la respuesta es correcta pero no se argumenta correctamente) o dos puntos (si la respuesta y su justificación son correctas) y reporta un alfa de Cronbach de 0.632 (Pensamiento Crítico, 2021). El costo de la prueba depende de la cantidad de

sustentantes, aproximadamente para 200 estudiantes, el costo es de 1,800 euros (información obtenida por correo electrónico).

El Watson Glaser Critical Thinking Appraisal es un examen de opción múltiple que consta de 40 reactivos de opción múltiple en los que se evalúa la habilidad para reconocer supuestos, evaluar argumentos y plantear conclusiones lógicas. Tiene una duración aproximada de 35-50 minutos, se aplica desde estudiantes que hayan terminado la secundaria hasta aquellos que se han graduado y los clasifica en nivel bajo, medio y superior (Pearson, 2021). Baweens y Gerhard lo utilizaron en estudiantes de enfermería en búsqueda de un instrumento predictivo de culminación exitosa, sin embargo, no encontraron relación significativa (Bauwens & Gerhard, 1987).

Algunos autores han utilizado los instrumentos ya mencionados y otros más como el Critical Thinking Assessment Test para evaluar el pensamiento crítico en estudiantes de licenciaturas del área de Ciencias; sin embargo, no son herramientas que evalúen el pensamiento crítico de áreas específicas (Haynes, Lisic, Goltz, Stein, & Harris, 2016; Styers, Van Zandt, & Hayden, 2018; Cox & McLaughlin, 2013).

Método para evaluar el pensamiento crítico con base en casos clínicos

Como se mencionó anteriormente, la mayoría de los instrumentos para la evaluación del pensamiento crítico se enfocan en situaciones de la vida diaria y no en áreas específicas. Por lo anterior, este trabajo se basa en una herramienta que originalmente evalúa el razonamiento clínico, que, como ya mencionamos, en este trabajo se tomará como el pensamiento crítico aplicado a una situación clínica (Noll, Key, & Jensen, 2001).

En 2010 el Centro Universitario de Utrecht desarrolló un formato de evaluación escrita basada en casos clínicos, el examen de razonamiento clínico basado en casos (ten Cate, 2018). Si bien, anteriormente se utilizaban ya algunas herramientas para la evaluación del razonamiento clínico como exámenes de opción múltiple (tres o cuatro), preguntas de emparejamiento extendidas (más de cinco opciones de respuesta), ensayos de respuesta corta o larga, problemas de manejo de pacientes (lista de selección de acciones a tomar), entre otras, todas con el común denominador de un escenario clínico (Daniel, et. al., 2019), este método une varias de ellas en un solo instrumento (ten Cate, 2018).

Inicialmente se presenta la información que tiene un médico al empezar la consulta, como edad, sexo y motivo de consulta y posteriormente se presentan preguntas

acerca del caso. En este formato se incluyen listas de opciones como en la relación de columnas, enfoques de diagnóstico diferencial como en los rompecabezas integrados e información adicional para alterar hipótesis como en los guiones de concordancia (ten Cate, 2018).

La característica principal de los CBCRT es que se presenta información inicial del paciente y después una lista con múltiples opciones de respuesta (5-25), incluyendo diagnósticos, hipótesis, datos de la exploración física, opciones de tratamiento, etc. Del caso presentado al inicio se derivan múltiples escenarios (generalmente 3), presentando información adicional del mismo paciente que puede cambiar la hipótesis inicial del estudiante. Todas las opciones son idénticas durante todo el caso, pero las opciones correctas varían dependiendo de la pregunta. Cada pregunta solicitará la selección de las respuestas que el estudiante considere pertinentes, es decir, no se limita a una sola respuesta, sino que da la oportunidad de marcar varias opciones que considere correctas. Para la puntuación del examen se toma como unidad cada opción. Si una pregunta tiene 4 posibles respuestas, cada una tiene un valor de 1 punto. Este tipo de exámenes se han aplicado desde el año 2010 y se ha encontrado un alfa de Cronbach de 0.73 (ten Cate, 2018).

Dentro de las limitaciones de este tipo de examen se encuentra la posibilidad de encontrar pistas en las preguntas (ya que la respuesta se selecciona de una lista), no dar opción a respuestas alternas al no ser un examen de construcción de respuesta y no permitir el ordenamiento jerárquico de los posibles diagnósticos o tratamientos (ten Cate, 2018).

Los autores proponen algunas sugerencias para los elaboradores de preguntas de este tipo de evaluación (ten Cate, 2018):

- Incluir en la presentación inicial la edad, el sexo y signo o síntoma principal.
- Siempre referirse al paciente específicamente y asegurarse de que esa pregunta requiera la información dada anteriormente, de no ser así, es probable que no sea una pregunta relativa al paciente.
- Presentar la información como normalmente lo haría un paciente real y no como una nota clínica.
- Ser específico con el número de respuestas requeridas para cada pregunta.
- Asegurarse de que la lista no contenga opciones repetidas, mutuamente excluyentes.
- Otorgar información objetiva (tiempo de evolución específico, edad real y no aproximada).
- Presentar información adicional seguida de una pregunta.

- Al presentar un nuevo escenario, iniciar con “suponga que...”
- Mantener independencia en cada pregunta, es decir, que la respuesta a una pregunta no dependa de la respuesta de la pregunta anterior.
- Tener un balance entre las preguntas de diagnóstico, exploración física, exámenes de laboratorio y tratamiento.
- Asegurarse de que otro experto revise las preguntas sin mostrarle las respuestas. Esto podría ayudar a considerar más respuestas de las que se tenían contempladas inicialmente.
- Presentar casos clínicos que reflejen situaciones médicas comunes.

Kim y colaboradores proponen 5 cualidades para que los casos clínicos tengan mayor efectividad:

- Coherentes con el nivel del estudiante
- Realistas
- Atractivos
- Desafiantes
- Sustentados en el conocimiento previo

(Kim, Brock, Phillips, & Keary, 2006)

Durante el primer año de la licenciatura, los estudiantes aún no están familiarizados con muchos términos médicos, ni son capaces de proponer diagnósticos diferenciales, por lo cual, este tipo de exámenes deben adaptarse a su nivel.

CAPÍTULO 4: PLANTEAMIENTO DEL PROBLEMA, JUSTIFICACIÓN, PREGUNTA DE INVESTIGACIÓN Y OBJETIVOS

Planteamiento del problema

En los últimos años se ha visto que los médicos egresados tienen dificultades para aplicar sus conocimientos y llegar a un diagnóstico clínico correcto, y muchas veces lo hacen con base en patrones conocidos que tienen disponibles porque tuvieron práctica en estos. Lo anterior se debe a que la memorización de algoritmos no es suficiente para el razonamiento clínico (Villarroel & Bernal, 2014).

Actualmente, en la Facultad de Medicina se cuenta con pocas herramientas de evaluación que permitan conocer el desarrollo de las competencias propuestas en el plan de estudios 2010. En el caso de la asignatura de Bioquímica y Biología Molecular, se busca contribuir con el desarrollo de cinco de las ocho competencias. Por tal motivo, se han propuesto estrategias de aprendizaje y herramientas de evaluación para facilitar la integración de la asignatura con la clínica (Cea,

Hernández, Salazar, Soto, & Matuz, 2019). A pesar de los esfuerzos realizados, aún no existen instrumentos que cuenten con las evidencias de validez suficientes para evaluar las competencias relacionadas con la asignatura de Bioquímica y Biología Molecular (ByBM).

Si bien el currículo de la Facultad incluye ciertas competencias, como el pensamiento crítico, la experiencia muestra que éste no se ha desarrollado y esto afecta directamente a los pacientes y a la sociedad en la que los médicos se desenvuelven, por lo que es necesario desarrollar técnicas que permitan hacer un diagnóstico del estado actual de esta competencia.

Justificación

En la educación basada en métodos tradicionales, el estudiante está acostumbrado a memorizar información sin realmente comprenderla; aunado a esto, la aplicación de métodos de evaluación para este tipo de aprendizaje ha traído como consecuencia un desarrollo mínimo en la capacidad de pensamiento crítico por parte del estudiantado. Debido a ello, en este trabajo se propone la creación de un método de evaluación que ayude a medir el pensamiento crítico y, de esta manera, conocer las áreas de oportunidad de los estudiantes e implementar nuevas estrategias de enseñanza.

Pregunta de investigación

- ¿Cuáles son las fuentes de evidencia de validez de una herramienta para la evaluación del pensamiento crítico en los estudiantes de la asignatura de Bioquímica y Biología Molecular?

Objetivo general:

- Diseñar una herramienta para evaluar el pensamiento crítico en los estudiantes de la asignatura de Bioquímica y Biología Molecular.

Objetivos específicos:

- Diseñar un caso clínico orientado a diabetes tipo 2 (DT2).
- Identificar las fuentes de evidencia de validez de la herramienta que se propone desarrollar.

CAPÍTULO 5: MÉTODO

Método

Diseño de investigación

Se trata de un estudio de desarrollo, cuantitativo, descriptivo, transversal (Hoyos, 2013).

Procedimiento

Se convocó a un grupo de expertos para el desarrollo de dos herramientas, el examen de pensamiento crítico y un examen de conocimientos generales para obtener evidencias de validez de relación otras variables, ya que, a pesar de que existe un examen departamental que se aplica a todos los estudiantes, la información contenida en este es confidencial.

Para lo anterior, se definió juicio de experto como *“una opinión informada de personas con trayectoria en el tema, que son reconocidas por otros como expertos cualificados en éste, y que pueden dar información, evidencias, juicios y valoraciones”* (Galicía, Balderrama, & Edel, 2017) y se propusieron los siguientes criterios de inclusión: 1) cinco o más años de experiencia como profesor de la asignatura de Bioquímica y Biología Molecular. 2) Experiencia en la elaboración de tablas de especificaciones y de reactivos.

Con base en estos criterios se formaron cuatro órganos colegiados (Tabla 1) con las siguientes funciones (INEE, 2017):

- Consejo Rector (CR): avaló el marco conceptual, el objeto de medida y el establecimiento de los puntos de corte y propuso a los integrantes de los demás cuerpos colegiados. Se conformó por tres médicos cirujanos y una bióloga.
- Comité de Diseño (CD): su función fue seleccionar, delimitar y justificar el contenido de los instrumentos de evaluación. En este órgano participaron dos médicos cirujanos, dos biólogos y un químico farmacéutico biólogo.
- Comité de elaboración de reactivos (CER): fue el encargado de elaborar los reactivos del examen de conocimientos generales (ECG). A este comité se agregó como criterio de selección que los profesores impartieran el bloque del que desarrollarían los reactivos y estuvo formado por dos médicos cirujanos, cuatro biólogos y un químico farmacéutico biólogo.
- Comité de Validación 1 (CV1): su función fue seleccionar los reactivos para conformar ambos instrumentos, verificar que los mismos estuvieran alineados con los resultados de aprendizaje propuestos y fueran apropiados para el nivel académico de los estudiantes. En este comité participaron cuatro médicos cirujanos, dos biólogos y un químico farmacéutico biólogo.

Posteriormente se invitó a otro grupo de expertos que fueran profesores de ByBM y/o médicos con 3 o más años de experiencia en el tratamiento de pacientes con DT2 puesto que este fue el escenario clínico elegido por el CR para elaborar el instrumento para evaluar el pensamiento crítico. Con este grupo de expertos se formó un quinto cuerpo colegiado:

- Comité de validación 2 (CV2): se encargó de verificar que los reactivos del examen de pensamiento crítico (EPC) estuvieran alineados con los resultados de aprendizaje propuestos y proponer modificaciones para mejorar el instrumento. En este comité participaron dos biólogos, un químico farmacéutico biólogo, tres médicos cirujanos, tres médicos internistas, un médico urgenciólogo, siete médicos residentes de medicina familiar, un médico residente de nefrología y tres médicos residentes de medicina interna.

Tabla 1. Cuerpos colegiados

Nombre del comité o consejo	Criterios de inclusión	Participantes
Rector	<p>Contar con licenciatura en Medicina, Química, Biología o Químico Farmacéutico Biólogo.</p> <p>Experiencia en docencia (mínimo 5 años) y elaboración de reactivos.</p> <p>Conocimientos del programa académico de Bioquímica.</p> <p>Conocimientos de Bioquímica y Biología Molecular.</p>	<p>DMM.B.B2</p> <p>JPPV.MG</p> <p>KDLM.MG.B2</p> <p>OBT.MG</p>

	Ser parte del cuerpo docente del departamento de Bioquímica y Biología Molecular.	
Diseño	<p>Contar con licenciatura en Medicina, Química, Biología o Químico Farmacéutico Biólogo.</p> <p>Experiencia en docencia (mínimo 5 años) y elaboración de reactivos.</p> <p>Conocimientos del programa académico de Bioquímica.</p> <p>Conocimientos de Bioquímica y Biología Molecular.</p> <p>Ser parte del cuerpo docente del departamento de Bioquímica y Biología Molecular.</p> <p>Impartir el bloque asignado.</p>	<p>OILB.QFB.B1</p> <p>HVM.B.B1</p> <p>DMM.B.B2</p> <p>KLM.MG.B2</p> <p>DRHE.MG.B3</p> <p>HVM.B.B3</p>
Elaboración de reactivos	<p>Contar con licenciatura en Medicina, Química, Biología o Químico Farmacéutico Biólogo.</p> <p>Experiencia en docencia (mínimo 5 años) y elaboración de reactivos.</p> <p>Conocimientos del programa académico de Bioquímica.</p> <p>Conocimientos de Bioquímica y Biología Molecular.</p> <p>Ser parte del cuerpo docente del departamento de Bioquímica.</p> <p>Impartir el bloque asignado.</p>	<p>OILB.QFB.B1</p> <p>JVTV.B.B1</p> <p>REMC.B.B1</p> <p>MTEG.B.B2</p> <p>DMM.B.B2</p> <p>KLM.MG.B2</p> <p>HVM.B.B3</p> <p>DRHE.MG.B3</p> <p>ELH.B.B3</p>
Validación 1	<p>Contar con licenciatura en Medicina o Química o Biología o Químico Farmacéutico Biólogo.</p> <p>Experiencia en docencia (mínimo 5 años) y elaboración de reactivos.</p>	<p>OILB.QFB.B1</p> <p>HVM.B.B1</p> <p>DMM.B.B2</p>

	<p>Conocimientos del programa académico de Bioquímica.</p> <p>Conocimientos de Bioquímica y Biología Molecular.</p> <p>Ser parte del cuerpo docente del departamento de Bioquímica.</p>	<p>KLM.MG.B2</p> <p>DRHH.MG.B3</p> <p>OBT.MG</p> <p>JPP.MG</p>
Validación 2	<p>Perfil profesional (formación): Contar con licenciatura en Química, Biología, Químico Farmacéutico Biólogo o Medicina General o ser médico especialista o residente de Medicina Interna, de Urgencias, Familiar o Nefrología.</p> <p>Experiencia en docencia (mínimo 5 años) y elaboración de reactivos y/o tres o más años de experiencia en el tratamiento de pacientes con diabetes tipo 2.</p>	<p>JAPL.MIR2</p> <p>APE.MIR2</p> <p>JMZV.MIR4</p> <p>YMG.MFR1</p> <p>KSMD.MFR1</p> <p>AHV.MFR1</p> <p>ANB.MFR1</p> <p>MCG.MFR1</p> <p>ASR.MFR1</p> <p>EMG.MFR1</p> <p>ILP.MNR4</p> <p>HICT.MU</p> <p>OILB.QFB.B1</p> <p>HVM.B.B1</p> <p>DMM.B.B2</p> <p>KLM.MG.B2</p> <p>DRHE.MG.B3</p> <p>JPP.MG</p> <p>ECS.MI</p> <p>GHP.MI</p> <p>JOS.MI</p>

Con base en los 12 pasos del método de Downing y Haladyna (Downing S., 2006) el CR estableció el plan general para el desarrollo de las herramientas (Anexo 1 y 2) y determinó que la selección del contenido de los instrumentos se haría con base en la importancia de cada tema para el cumplimiento de los objetivos de la asignatura (OA) y para la comprensión de la DT2, tomando en cuenta únicamente los tres primeros bloques —agua, metabolismo de carbohidratos y lípidos—:

Objetivos generales de la asignatura:

1. Conocer los mecanismos moleculares del funcionamiento del organismo humano en el estado de salud y de enfermedad.
2. Integrar el conocimiento de los procesos biológicos moleculares en la estructura y función de las células, tejidos y el organismo.
3. Utilizar los conocimientos bioquímicos y de biología molecular como herramientas para el diagnóstico de las enfermedades.

La clasificación se estableció de la siguiente manera:

Sin importancia (SI)= el dominio de este tema no ayuda a cumplir los objetivos de la asignatura

Poco importante (PI)= el dominio de este tema favorece el cumplimiento de uno de los objetivos.

Importante (I)= el dominio de este tema es indispensable para cumplir con 2 o 3 objetivos de la asignatura.

La importancia para la comprensión de la DT2 se estableció de la siguiente manera:

Sin importancia (SI)= el dominio de este tema no ayuda a comprender la DT2.

Poco importante (PI)= favorece la comprensión de la DT2, sin embargo, el estudiante podría comprenderla sin dominar este tema.

Importante (I)= es indispensable el dominio de este tema para comprender la DT2.

Con un correo electrónico se envió a cada experto del CD una tabla del bloque correspondiente para su llenado con base en su experiencia. Esta tabla incluía *unidad, tema, subtema y objetivo de aprendizaje; el número de horas dedicado para su enseñanza; importancia del objetivo de aprendizaje para el cumplimiento de los objetivos de la asignatura (SI, PI, I); la importancia del objetivo para la comprensión de la DT2 (SI, PI, I); el número de reactivos propuesto en cada caso; el nivel taxonómico propuesto; y, el resultado de aprendizaje (en caso de ser importante para la comprensión de la DT2) (Tabla 2).*

Tabla 2. Tabla de especificaciones

Asignatura: Bioquímica y Biología Molecular												
UNIDAD	Tema	Subtema	Objetivo	HORAS	Importancia para el cumplimiento de los objetivos de la asignatura (OA)	Importancia para la comprensión de la diabetes tipo 2 (DT2)	# de reactivos		Nivel taxonómico		Resultado de aprendizaje	
							OA	DT2	OA	DT2	OA	DT2
2	1 Fundamentos del metabolismo celular	1.1 Vía metabólica	1.1.1 Discutir el concepto de vía metabólica									

Del mismo modo, junto con la tabla, se envió una guía con instrucciones para llenar la tabla de forma correcta que incluía un resumen de los niveles taxonómicos y los verbos recomendados para cada uno de ellos; así como recomendaciones para su redacción y una lista de verificación. Por su parte, los académicos llevaron a cabo sus propuestas de forma individual y llenaron los campos faltantes de acuerdo con su experiencia.

Una vez que los profesores enviaron por correo electrónico las tablas con sus respuestas, se convocó al CR y al CD a sesiones presenciales. En estas sesiones se revisaron minuciosamente las tablas y, teniendo en cuenta los contenidos del programa, los comités concluyeron que de los 162 objetivos de aprendizaje propuestos en los primeros tres bloques eran necesarios únicamente 66 para la comprensión de la DT2 y en total se propusieron 150 resultados de aprendizaje. Estos se dividieron en ocho partes - 2 de primer bloque, 3 de segundo bloque y 3 de tercer bloque - y se enviaron a los miembros del CER a quienes se les solicitó elaborar dos reactivos de opción múltiple (con 3 opciones de respuesta) por cada resultado de aprendizaje para elaborar un examen de conocimiento generales (ECG) con la intención de contar con una herramienta para obtener evidencias de validez de relación con otras variables.

Posteriormente se convocó a una reunión con el CD, se seleccionaron 44 reactivos; se enviaron junto con un cuestionario (Jiménez, Flores, Alpuche, & Martínez, 2016) en google forms 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 que consistió en 34 reactivos con tres opciones de respuesta cada uno.

En función de los resultados de aprendizaje el ECG, los reactivos y los atributos a evaluar del pensamiento crítico, se propusieron los resultados de aprendizaje (relacionados con el contenido del examen de opción múltiple) para el EPC y posteriormente los reactivos con base en el método propuesto por Ollen ten Cate - exámenes de razonamiento clínico basados en casos, CBCRT por sus siglas en inglés (ten Cate & Loon, 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 una o más respuestas de una lista que incluye 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” y “hormonas y neurotransmisores”.

Una vez aprobadas ambas herramientas por los comités correspondientes, se convocaron seis jueces para establecer el punto de corte. Se les envió por correo electrónico un artículo sobre el método y se convocó a una reunión en línea vía zoom para explicar en qué consistía el método. Posteriormente se llevó a cabo la reunión para establecer el punto de corte (Tabla 4, 5 y 6). (Canadas's Testing Company, 2014).

Finalmente se invitó a 41 estudiantes (muestreo no probabilístico por conveniencia) (Hernández, 2014; Pimienta, 2000) 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. A los estudiantes se les envió previo al examen el consentimiento informado, aviso de privacidad, acuerdo de confidencialidad y una carta de invitación. Para facilitar el proceso, no se les pidió que entregaran los documentos firmados, únicamente se agregaron preguntas en el examen en las que aceptaban que contaban con los documentos y aceptaban participar.

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 para la herramienta de pensamiento crítico, se resolvió un caso clínico de prueba proyectando la pantalla.

Para ambos instrumentos se calculó promedio, mediana, desviación estándar, grado de dificultad, índice de discriminación, punto biserial y alfa de Cronbach.

CAPÍTULO 6: RETOS ANTE LA PANDEMIA POR SARS-COV-2

Al encontrarnos ante una situación de emergencia sanitaria debido a la aparición del Covid-19 (Cortés & Ponciano, 2021), la investigación sufrió ciertos retrasos, puesto que el personal de salud, en donde se incluían la mayoría de los expertos, no contaban con el tiempo suficiente para llevar a cabo las tareas requeridas para continuar en tiempo y forma este proyecto. Fue así, que, inicialmente pese a que la mayoría de los expertos aceptaron la invitación, únicamente se recibió la documentación en tiempo y forma por parte de cuatro docentes, dos médicos cirujanos, dos médicos residentes de Medicina Interna y dos especialistas en Medicina Interna y docentes. Sin embargo, del total de invitados que entregaron la documentación, solamente los docentes integrantes del CV1 y tres médicos especialistas en Medicina Interna revisaron la información por lo que hubo necesidad de invitar a otros expertos, con los que finalmente se conformaron los comités correspondientes.

Por otro lado, a pesar de que las herramientas de evaluación estuvieron listas para ser aplicadas en tiempo y forma, hubo poca participación por parte de los estudiantes ya que, al estar inconformes con la universidad por cuestiones administrativas, hubo suspensión de actividades dos veces (La Jornada, 2021),

aproximadamente dos semanas en cada ocasión, motivo por el cual se retrasó el curso y se recorrieron las fechas de exámenes. De tal modo, en las fechas en las que fue programada la aplicación de las herramientas de evaluación, únicamente asistieron 30 de 300 estudiantes invitados para participar en la prueba piloto y no hubo oportunidad de aplicarlo de forma definitiva como era la intención inicial para tener resultados sobre el nivel de pensamiento crítico de los estudiantes.

CAPÍTULO 7: RESULTADOS

A continuación, se presentará la estructura de los instrumentos mas no el instrumento ya que sigue pendiente la aplicación del instrumento en un grupo mayor y mostrarlo en este trabajo sería una amenaza a la validez.

Para la evaluación de los 44 reactivos seleccionados del examen de conocimientos generales se envió un cuestionario a seis jueces expertos integrantes de los comités.

El cuestionario contenía 14 dominios a evaluar y una sección para comentarios y sugerencias sobre cada reactivo. Con base en la opinión de los expertos se seleccionaron 34 reactivos ya que 10 de ellos evaluaban resultados de aprendizaje semejantes. Es importante mencionar que algunos jueces eran parte del CER por lo que no se tomaron en cuenta las respuestas de aquellos que eran autores de las preguntas. Como se puede ver en la Anexo 2, en la primera revisión no había total acuerdo en cuanto a las características de ciertos reactivos por lo que se convocó a una reunión en línea mediante zoom para hacer las modificaciones correspondientes. Una vez que hubo total acuerdo entre los jueces, se seleccionaron los 34 reactivos para conformar la versión final del instrumento.

Con base en este instrumento, se diseñó el examen del EPC. Esto se hizo con base en 24 resultados de aprendizaje relacionados con los anteriormente redactados en el

ECG. A partir de ellos se plantearon 7 preguntas con 24 respuestas en total. Lo anterior, debido a que, cada pregunta puede tener más de una solución.

La revisión de la herramienta se dividió en tres fases. En la primera se convocó a seis miembros del CV2; con base en sus observaciones se hicieron modificaciones y se envió a tres miembros más. Posteriormente se envió a los médicos residentes y un médico especialista en urgencias para que respondieran la herramienta con la intención de que su acercamiento fuera como el de los estudiantes y después se les envió el material completo para que hicieran sus observaciones. En la tabla se muestra el cambio de la estructura del examen después de cada revisión.

Tabla 3. Evolución de la estructura del examen del pensamiento crítico

Presentación de Caso clínico					
	Número de respuestas				
Escenario A	Versión inicial	Primera revisión	Segunda revisión	Tercera revisión	Versión final
Pregunta 1	3	3	4	4	4
Pregunta 2	2	2	2	2	2
Pregunta 3	7	7	7	7	7
Escenario B					
Pregunta 4	1	1	2	2	2
Pregunta 5	1	1	1	1	1
Escenario C					
Pregunta 6	3	3	3	3	3
Pregunta 7	3	3	3	4	5
Opciones de respuesta					
Columna 1	11	11	11	11	11

Columna 2	12	12	12	12	12
Columna 3	21	21	21	21	21
Columna 4	17	17	17	17	17

Después de tener las versiones finales de ambos instrumentos, se establecieron los puntos de corte con base en el método de Angoff. Para el ECG se definió al estudiante mínimamente competente como aquel que *sabe leer, comprende los conceptos generales de Bioquímica, la semántica en la asignatura y lo aplica en la solución de problema* y se estableció el punto de corte en 75.4. (Tabla 4).

Tabla 4. Probabilidad (%) de que un estudiante mínimamente competente responda correctamente los reactivos del examen general de conocimientos.

Número de Reactivo	Juez 1	Juez 2	Juez 3	Juez 4	Juez 5	Promedio
1	50	70	50	50	50	54
2	70	90	70	80	70	76
3	40		50		50	46.7
4	50	70	50	70	60	60
5	60	50	60	65	60	59
6	40	70	70	70	50	60
7	70	70	80	70	60	70
8	80	80	80	80	80	80
9	40	60	60	60	50	54
10	80	70	80	80	80	78
11	100	90	90	90	95	93
12	100	90	90	90	95	93
13	80	60	80	60	80	72
14	80	90	80	80	95	85
15	90	90	70	70	70	78

16	90	90	70	70	85	81
17	80	90	70		70	77.5
18	90	70	70	70	90	78
19	90	90	70	75	85	82
20	90	90	80	80	90	86
21	90	90	80	80	90	86
22	90	70	80	80	90	82
23	100	100	90	85	95	94
24	90	90	80	80	90	86
25	90	90	70	65	80	79
26	90	90	80	90	90	88
27	70	70	80	90	80	78
28	70	80	80	80	80	78
29	100	90	90	75	90	89
30	90	90	90	90	90	90
31	90	90	90	75	90	87
32	70	70	70	85	80	75
33	90	90	90	80	90	88
34	90	90	90	90	90	90
Punto de corte para el examen de conocimientos generales						75.4

Para el examen de pensamiento crítico se establecieron dos niveles y sus definiciones se hicieron con base en las competencias que se evalúan en el instrumento. El estudiante mínimamente competente para el nivel uno se definió como aquel que *identifica, selecciona e interpreta información para plantear y/o resolver un problema* y el punto de corte se estableció en 72.9 (Tabla 5).

Tabla 5. Probabilidad (%) de que un estudiante mínimamente competente del nivel 1 responda correctamente los reactivos del examen del pensamiento crítico.

Número de reactivo	Juez 1	Juez 2	Juez 3	Juez 4	Juez 5	Juez 6	Promedio
1	80	80	70	75	80	85	78.
2	80	80	70	75	80	85	78.3
3	80	80	70	75	80	85	78.3
4	70	75	75	70	70	85	74.2
5	60	70	70	70	70	80	70
6	60	60	75	70	65	60	65
7	70	70	60	60	65	75	66.7
8	60	70	60	60	65	75	65
9	70	60	70	65	75	70	68.3
10	70	60	70	65	75	70	68.3
11	80	60	70	60	75	70	69.2
12	70	60	70	60	70	60	65
13	60	60	70	75	70	60	65.8
14	90	80	75	75	80	80	80
15	60	60	70	65	80	60	65.8
16	70	70	70	70	70	70	70
17	90	80	80	80	85	90	84.2
18	80	80	80	80	80	80	80
19	90	90	80	85	90	80	85.8
20	80	80	80	80	80	80	80
21	80	80	80	80	80	70	78.3
22	70		70	70	70	70	68.3
23	70	75	70	75	75	80	74.1
24	70	75	70	70	70	75	71.6
Punto de corte para el nivel 1 del examen de pensamiento crítico							72.9

El estudiante mínimamente competente para el nivel dos se definió como aquel que *identifica, selecciona e interpreta información para plantear y/o resolver un problema y analiza y discierne información para desarrollar el pensamiento crítico*. Para este nivel se estableció el punto de corte en 85.7 (Tabla 6).

Tabla 6. Probabilidad (%) de que un estudiante mínimamente competente del nivel 2 responda correctamente los reactivos del examen del pensamiento crítico.

Número de reactivo	Juez 1	Juez 2	Juez 3	Juez 4	Juez 5	Juez 6	Promedio
1	80	85	80	75	85	85	81.7
2	80	80	80	75	85	85	80.8
3	80	80	80	75	85	85	80.8
4	80	80	80	70	80	80	78.3
5	80	90	80	85	90	75	83.3
6	80	85	85	75	90	75	81.7
7	90	95	80	80	85	90	86.7
8	90	95	80	80	85	90	86.7
9	85	90	85	80	95	85	86.7
10	85	90	85	80	95	85	86.7
11	90	90	85	80	95	85	87.5
12	85	90	85	80	90	85	85.8
13	85	90	85	90	90	85	87.5
14	90	90	90	85	90	90	89.2
15	80	85	85	85	85	90	85.0
16	90	95	85	85	90	90	89.2
17	90	80	85	85	90	90	86.7
18	90	85	85	85	90	90	87.5
19	90	90	85	95	90	90	90.0
20	90	90	85	90	90	85	88.3
21	90	90	85	90	90	85	88.3
22	80	90	85	85	90	85	85.8
23	90	85	80	90	80	90	85.8
24	85	90	90	80	90	80	85.8
Punto de corte para el nivel 2 del examen de pensamiento crítico							85.7

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 (Mártinez, Hernández, & Hernández, 2014a) utilizando Excel. Obtuvimos para 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) (Chávez & Saade, 2010) (Anexo 3), varianza y alfa de Cronbach (Barrios & Cosculluela, 2013) y se establecieron como útiles y seleccionables aquellos que tuvieron grado de dificultad (GD) > .20, índice de discriminación (ID) >.15 y punto biserial >.15.

Los resultados psicométricos mostraron un alfa de Cronbach de 0.83 para el examen de conocimientos generales y de 0.46 para el examen de pensamiento crítico y únicamente en este último se alcanzó el punto de corte como se muestra en la Tabla 7.

Tabla 7. Comparativa de los datos estadísticos del examen de conocimientos generales y el de pensamiento crítico.

	Examen de conocimientos generales	Examen de pensamiento crítico
Promedio	64.9	49.1
Mediana	61.7	45.8
Calificación más baja	38.2	29.2

Calificación más alta	92	70.8
Desviación estándar	15.17	12.36
Grado de dificultad	64.9	49.1
alfa de Cronbach	.83	.46*

Se identificaron 12 reactivos (5 del examen de conocimientos generales y 7 del examen de pensamiento crítico) que es necesario revisar (Tabla 8 y 9), ya que no cumplieron con los criterios estadísticos establecidos por el CV para ser seleccionados como útiles: $GD > .20$, $ID > .15$, punto biserial $> .15$.

Tabla 8. Reactivos del examen general de conocimientos que no cumplieron con los criterios estadísticos establecidos por el CV.

Examen de conocimiento generales			
Número de reactivo	Grado de dificultad	Índice de discriminación	Punto biserial
1	53.3	-0.22	-0.2
13	73.3	0.00	0.11
23	90.0	-0.11	-0.15
27	23.3	0.11	0.09
32	93.33	0.11	0.16

Tabla 9. Reactivos del examen de pensamiento crítico que no cumplieron con los criterios estadísticos establecidos por el CV.

Examen de pensamiento crítico			
Número de reactivo	Grado de dificultad	Índice de discriminación	Punto biserial
1	82.8	0.13	0.28

2	31.03	0.08	-0.03
9	68.97	0.14	0.23
10	72.41	0.03	0.11
12	27.59	0.08	0.12
16	6.9	0.11	0.21
23	24.34	0.08	0.15

La programación inicial al invitar a los estudiantes a participar en la aplicación de los exámenes fue de tres horas para tener tiempo extra en caso de cualquier contratiempo; 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 finalizar el examen se aplicó una encuesta con preguntas para conocer la experiencia de los estudiantes: ¿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 cómo 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 [...]”

CAPÍTULO 8: DISCUSIÓN

Como propuso Van der Vleuten, para determinar la utilidad de un método de evaluación, en el diseño de este instrumento se buscó que fuera confiable y aceptable tanto por profesores como por estudiantes y que tuviera evidencias de validez e impacto en el aprendizaje. (Epstein, 2007; Downing & Haladyna, 2009b).

Si bien es cierto que los exámenes de opción múltiple han recibido fuertes críticas a su efectividad por tener limitaciones como la probabilidad de respuesta al azar, la aparición de señales sobre la respuesta correcta dentro de la formulación de la pregunta y la presencia de contenido trivial que únicamente evalúe la capacidad de recordar, (Downing, 2002; Downing & Haladyna, 2009b) con este instrumento se disminuyen considerablemente gracias a la cantidad de respuestas que existen y la evaluación por múltiples expertos.

Con base en las recomendaciones de Downing y Haladyna (Downing S., 2006) seguimos el proceso para obtener evidencias de validez de contenido, del proceso de respuesta, de estructura interna, de relación con otras variables y de consecuencias (Downing & Haladyna, 2009b).

En el caso de la validez de contenido se contó con un grupo de expertos que con base en su conocimiento y expertise seleccionaron el contenido para evaluar el constructo (Lozano & Turbany, 2013) y determinaron que los reactivos incluidos eran idóneos,

precisos y representativos. Tal como lo sugiere Berk (1990), la selección del contenido se hizo inicialmente de manera individual y posteriormente se llevó a cabo una discusión en grupo (Berk, 1990).

Para asegurar las evidencias del proceso de respuesta y evitar tener varianza irrelevante de constructo por falta de familiaridad con el instrumento, antes de iniciar el examen se explicó en qué consistía la herramienta y se mostró y resolvió un caso clínico con los estudiantes. Del mismo modo, al finalizar se hicieron preguntas a los estudiantes sobre el formato del examen como lo sugieren Padilla y Benitez (Padilla & Benítez, 2014). Sin embargo, a pesar de lo anterior, es posible que el alfa de Cronbach obtenida sea resultado del desconocimiento de la estrategia.

Como evidencias de validez para la estructura interna se realizaron las pruebas psicométricas con base en la teoría clásica y se calculó el alfa de Cronbach (Chávez & Saade, 2010). Para evidenciar la relación con otras variables se elaboró un examen de conocimientos generales y a pesar de que el examen no fue de altas consecuencias, sino que se propuso para hacer un diagnóstico del pensamiento crítico en los estudiantes, se estableció el punto de corte con base en el método de Angoff (Canadas's Testing Company, 2014).

Con relación a la confiabilidad, ten Cate reportó un alfa de Cronbach de 0.73 en los exámenes de razonamiento clínico basado en casos (método que se utilizó para

diseñar el EPC) (ten Cate, 2018); sin embargo, es importante mencionar que su experiencia es mayor ya que los estudiantes están relacionados con la herramienta pues en la universidad de Utrecht se ha aplicado por más de 10 años y en nuestro caso al ser calculada con base en la varianza y ser un examen que los estudiantes no conocían era esperado tener un valor menor al óptimo (Chávez & Saade, 2010). Por lo anterior, es necesario aplicar el instrumento en una población más grande para observar mejor el comportamiento.

Con respecto al tiempo, observamos lo mismo que lo reportado por ten Cate (ten Cate, *Assessment of Clinical Reasoning Using the CBCR Test*, 2018) ya que, en ambos casos, la duración de la prueba fue menor a una hora a pesar de que en el método original hay cuatro preguntas y en el nuestro hubo siete.

En el método original no se reporta el método para establecer el punto de corte, pero en este caso llama la atención que el únicamente el 30% de los estudiantes obtuvo una calificación mayor o igual al punto de corte en el ECG mientras que en el EPC ningún estudiante alcanzó el punto de corte del nivel uno ni el del nivel dos. Por el momento no podemos hacer inferencias sobre el nivel de pensamiento crítico de los estudiantes porque se trata de una prueba piloto por lo que atribuimos este dato a la falta de familiaridad con el instrumento y de la misma manera que con la confiabilidad, es necesario aplicarlo en una población mayor. Por otro lado, es

importante mencionar que en este caso, el juicio se emitió por parte de los expertos con conocimiento de la respuesta correcta, fenómeno que pudo haber causado que la calificación emitida fuera alta; por lo tanto que el punto de corte lo fuera también y que pocos estudiantes alcanzaran la calificación esperada en el ECG y ninguno en el EPC. Con respecto a lo anterior, Verheggen, et. al. reportan que el juicio emitido por expertos que conocen la respuesta a la pregunta y los que no, es diferente e influye en qué tan alto es el juicio que emiten (Verheggen, Muijtjens, Van Os, & Schuwirth, 2018).

Los reactivos que identificamos fuera de los parámetros establecidos se revisaron por el CR, CD y nuevamente por los CV y se hicieron las modificaciones pertinentes para su aplicación definitiva.

CAPÍTULO 9: CONCLUSIONES

A pesar de que existen múltiples instrumentos para la evaluación del pensamiento crítico, es importante definir cuál es el objetivo de medir éste, ya que las herramientas conocidas se centran en la evaluación de las habilidades del pensamiento crítico en general por lo que es necesario contar con herramientas que evalúen áreas específicas.

Con el reporte preliminar podemos concluir que es una herramienta con buena 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 y en este caso, la duración no fue mayor a 1 hora, lo cual coincide con lo reportado por ten Cate (ten Cate, 2018).

Hasta el momento no podemos hacer inferencias por tratarse de una prueba piloto; sin embargo, el proyecto continuará en marcha ya que la siguiente meta a alcanzar es aplicarlo en toda la generación, obtener información para conocer las áreas de oportunidad y de esta manera proponer estrategias para fomentar el pensamiento crítico en los estudiantes de Medicina.

LIMITACIONES Y PERSPECTIVAS

Para este proyecto únicamente se tomó en cuenta a estudiantes de la Facultad de Medicina de la UNAM y el programa de la asignatura de Bioquímica de la misma Institución.

Si bien uno de los principales objetivos de este trabajo era aplicar los fundamentos teóricos para lograr evaluar el pensamiento crítico en los estudiantes de medicina, las circunstancias no lo permitieron.

Sin embargo, una de las metas en la siguiente etapa de esta investigación es comparar el nivel del pensamiento crítico de estudiantes cuya formación sigue el sistema actual con aquellos cuya formación incluye el aprendizaje basado en problema.

Con los resultados obtenidos se propondrán estrategias educativas que fomenten el pensamiento crítico de los estudiantes.

REFERENCIAS

1. Anthony, G. (1996). Active Learning in a constructivist framework. En *Educational studies in Mathematics*, 349-389.
2. Arrekkuzhiy1, S. (2019). Assessment Practices in Higher Education: Myths and Realities. *University News*, 18-20.
3. Bailin, S., Case, R., Coombs, J., & Daniels, L. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies*, 285-302.
4. Barrios, M., & Cosculluela, A. (2013). Fiabilidad. En M. B. Meneses, *Psicometría*, 75-140. Barcelona: UOC.
5. Bauwens, E., & Gerhard, G. (1987). The Use of the Watson-Glaser Critical Thinking Appraisal to Predict Success in a Baccalaureate Nursing Program. *Journal of Nursing Education*, 278-281.
6. Berk, R. (1990). Importance of expert judgment in content-related validity evidence. *Western Journal of Curriculum Studies*, 650-671.
7. Bezanilla, M., Poblete, M., Fernández, D., Arranz, S., & Campo, L. (2018). El pensamiento crítico desde la perspectiva de los docentes universitarios. *Estudios Pedagógicos*, 89-113.
8. Bloom, B., Engelhart, M., Frust, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives*. 62-197 Michigan: Longmans.

9. Boshuizen, H., & Schmidt, H. (1992). On the role of Biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Sciences*, 153-184.
10. Butler, H. D., Butler, H., Dwyer, C., Hogan, M., Franco, A., Rivas, S., . . . Almeida, L. (2012). The Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity*, 112-121.
11. Calle, G. (2013). La evaluación de las habilidades del pensamiento crítico asociadas a la escritura digital. *Revista virtual universitaria Católica del Norte*, 68-83.
12. Canadas's Testing Company. (2014). *The Angoff Method Standard Setting*. Ottawa: *Assessment strategies*, 1-4.
13. Carrasco, 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.
14. Castañeda, M. (2015). La enseñanza-aprendizaje por competencia. *Revista de la Facultad de Derecho*, 243-265.
15. Cea, M., Hernández, D., Salazar, M., Soto, I., & Matuz, D. (2019). El uso de escenarios clínicos y el aprendizaje de Bioquímica en alumnos de primer año de la carrera de Medicina. *Revista de Investigación Médica*, 187-192.

16. Chávez, C., & Saade, A. (2010). Fundamentos técnicos del análisis de reactivos. En *Procedimientos básicos para el análisis de reactivos*, 49-97. Ciudad de México: Centro Nacional de Evaluación para la Educación superior.
17. Consejo Técnico de la Facultad de Medicina. (2010). *Plan de Estudios de la Licenciatura de Médico Cirujano 2010*. Ciudad de México.
18. Cortés, A., & Ponciano, G. (2021). Impacto de los determinantes sociales de la COVID-19 en México. *Boletín sobre COVID-19*, 9-13.
19. Cox, W., & McLaughlin, J. (2013). *Association of Health Sciences Reasoning Test Scores with Academic and Experiential Performance*. Carolina: *American Journal of Pharmaceutical Education*, 1-5.
20. Daniel, M., Rencic, J., Durning, S., Holmboe, E., Santen, S., Lang, V., . . . Gruppen, D. (2019). Clinical Reasoning Assessment Methods: A Scoping Review and Practical Guidance. *Academic Medicine*, 902-912.
21. DOF. (2017). *Criterios técnicos para el desarrollo, uso y mantenimiento de instrumentos de evaluación*, 1-37. México.
22. Downing, S. (2002). Assessment of knowledge with written test forms. En Norman, Van der Vleuten & Newble, *International Handbook of Research in Medical Education*, 647-672. Springer.

23. Downing, S. (2006). Twelve steps for effective test development. En Lane, Raymond & Haladyna, *Handbook of test development*, 3-25. New York and London: Routledge.
24. Downing, S. & Yudkowsky, R. (2009). Introduction to Assessment in the Health Professions. En Downing & Yodkowsky, *Assessment in Health Professions Education* (págs. 1-20). New York: Routledge.
25. Downing, S., & Haladyna, T. (2009). Validity and its threats. En Downing & Haladyna, *Assessment in Health Professions Education*, 21-55. New York: Routledge.
26. Durante, M., Lozano, J., Martínez, A., Morales, S., & Sánchez, M. (2012). Introducción. En M. L. Durante, *Evaluación de Competencias en Ciencias de la Salud*, 2-8. México: Editorial Médica Panamericana.
27. Epstein, R. (2007). Assessment in Medical Education. *The New England Journal of Medicine*, 387-396.
28. ETSU. (2020). *Testing Center, Center for academic Achievement*. Obtenido de <https://www.etsu.edu/students/cfaa/testing/test/cctst.php>
29. Facione, P. (1998). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*, 1-19.
30. Facione, P. (2016). Critical thinking: what it is and why it counts. *Measured reasons LLC*, 1-30.

31. Fajardo, G. (2019). *Informe Facultad de Medicina 2016-2020*. México.
32. Ferris, H., & Flynn, D. (2015). Assessment in Medical Education: What are we trying to achieve? *International Journal of Higher Education*, 139-144.
33. 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.
34. Galicia, L., Balderrama, J., & Edel, R. (2017). Validez de contenido por juicio de expertos: propuesta de una herramienta virtual. *Apertura*, 42-53.
35. González, J., Wagenaar, R., & Benetione, P. (2004). Tuning-América Latina: Un Proyecto de las Universidades. *Revista Iberoamericana de Educación*, 151-164.
36. Halpern, D. (1998). Teaching critical thinking for transfer across domains. . *American Psychologist*, 449-455.
37. Harasym, P., Tsai, T., & Hemmati, P. (2008). Current trends in developing medical students' critical thinking abilities. *Elsevier*, 341-355.
38. Haynes, A., Lisic, E., Goltz, M., Stein, B., & Harris, K. (2016). Moving Beyond Assessment to Improving Students' Critical Thinking Skills: A Model for

- Implementing Change. *Journal of the Scholarship of Teaching and Learning*, 44-61.
39. Hernández, R. (2014). Selección de muestra. En C. & Fernández, *Metodología de la investigación*, 170-194. México: McGraw Hill.
40. Hoyos, M. (2013). Estudios descriptivos. *Revista de Actualización Clínica Investiga*, 1670-1674.
41. INEE. (2017). *Criterios técnicos para el desarrollo, uso y mantenimiento de instrumentos de evaluación*. México.
42. Insight Assessment. (2021a). Obtenido de Tools for Health Science: <https://www.insightassessment.com/tools-for-health-science>
43. Insight Assessment. (2021b). Obtenido de Advancing Thinking Worldwide: <https://www.insightassessment.com>
44. Jiménez, J., Flores, F., Alpuche, A., & Martínez, A. (2016). Evaluación de reactivos de opción múltiple en Medicina. Evidencias de validez de un instrumento. *Revista de Investigación en Educación Médica*, 8-15.
45. Kaufman, J., & Reichert, K. (2018). "What is measured gets improved" (or if you cannot measure it, you cannot improve it)*. *Pediatric Critical Care Medicine*, 267-268.
46. Khan, M., & Aljarallah, B. (2011). Evaluation of modified questions (MEQ) and multiple choice questions (MCQ) as a tool for assessing the cognitive

- skills of undergraduate medical students. *International Journal of Health Sciences*, 39-43.
47. Kim, S., Brock, D., Phillips, K., & Keary, J. (2006). A conceptual framework for developing teaching cases: a review and synthesis of the literature across disciplines. *Medical Education*, 867-876.
48. Kuiper, R., & Pesut, D. (2004). Promoting cognitive and metacognitive reflective reasoning skills in nursing practice: self-regulated learning theory. *Issues and Innovation in nursing education*, 381-391.
49. Kunina, O., Hautz, W., Knigge, M., Spies, C., & Ahlers, O. (2015). Assessing clinical reasoning (ASCLIRE): Instrument development and validation. *Advances in Health Sciences Education*, 1205-1224.
50. La Jornada. (2021). *Quince facultades de la UNAM se suman al paro por retraso de pagos*. Obtenido de *Quince facultades de la UNAM se suman al paro por retraso de pagos*.
51. Lozano, L., & Turbany, J. (2013). Validez. En J. B. Meses, *Psicometría* (págs. 141-200). Barcelona: UOC.
52. Mafinejad, M., Soltani, S., Monajemi, A., Jalili, M., Soltani, A., & Rasouli, J. (2017). Use of Multiple-Response format test in the assessment of medical student's critical thinking ability. *Journal of Clinical & Diagnostic Research*, 10-13.

53. Makary, M., & Daniel, M. (2016). Medical error - the third leading cause of death in the US. *BMJ*, 1-5.
54. Martínez, M., Hernández, M., & Hernández, M. (2014,a). Teoría Clásica de los Tests (TCT) I: el modelo y la fiabilidad de las puntuaciones. En M. Martínez, M. Hernández, & M. Hernández, *Psicometría*, 37-99. Madrid: Alianza.
55. Martínez, M., Hernández, M., & Hernández, M. (2014,b). Introducción a la Teoría de la Generalizabilidad. En *Psicometría*, 101-122. Madrid: Alianza.
56. McDonald, P., Howard, S., Schlump, K., & Plack, M. (2014). Learning partnership: students and faculty learning together to facilitate reflection and higher order thinking in a blended course. *Online Learning Consortium*, 1-22.
57. Morales, J., Varela, M., & Núñez, A. (2015). Educación basada en competencias. En M. L. Sánchez, *Educación Médica: Teoría y Práctica*, 2-77. México: Elsevier.
58. Nieto, S., Saiz, C., & Orgaz, B. (2009). Análisis de la propiedades psicométricas de la versión española del HCTAES-Test de Halpern para la evaluación del pensamiento crítico mediante situaciones cotidianas. *Revista Electrónica de Metodología Aplicada*, 1-15.
59. Noll, E., Key, A., & Jensen, G. (2001). Clinical reasoning of an experienced physiotherapist: insight into clinician decision-making regarding low back pain. *Physiotherapy Research International*, 40-51.

60. Núñez, S., Ávila, J., & Olivares, S. (2017). El desarrollo del pensamiento crítico en estudiantes universitarios por medio del Aprendizaje Basado en Problemas. *Revista Iberoamericana de Educación Superior*, 84-103.
61. Padilla, J., & Benítez, I. (2014). Validity evidence based on response process. *Psicothema*, 136-144.
62. Paul, R., & Elder, L. (2003). La mini-guía para el pensamiento crítico, conceptos y herramientas. *Fundación para el pensamiento crítico*, 1-26.
63. Pearson. (2021). *Watson-Glaser Critical Thinking Appraisal*. Obtenido de https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/efficacy-and-research/reports/Watson-Glaser_One_Page_Summary.pdf
64. *Pensamiento Crítico*. (2021). Obtenido de <https://www.pensamiento-critico.com/evaluacion-test>
65. Pimienta, R. (2000). Encuestas probabilísticas vs no probabilísticas. *Política y Cultura*, 263-276.
66. RCSJ. (2021). *Health Science Reasoning Test – Associate Degree (HSRT-AD)*. Obtenido de Rowan College South Jersey: <https://www.rcsj.edu/testing/cumberland/hsrt-ad>
67. Riveros, H. (2015). Teaching to reason. *European J of Physics Education*, 54-59.

68. Rojas, G., & Díaz, F. (2012). Una mirada psicoeducativa al aprendizaje: qué sabemos y hacia dónde vamos. *Sinéctica Revista electrónica de educación*, 1-19.
69. Sadi, Ö. (2013). Case Study of How Turkish University Students Improve Their Biochemistry Achievement. *Higher Education Studies*, 52-67.
70. Sánchez, M. (2018). La evaluación del aprendizaje de los estudiantes: ¿es realmente tan complicada? *Revista Digital Universitaria*, 1-18.
71. Sánchez, M., & Martínez, A. (2014). Razonamiento clínico. En M. & Sánchez, *Informática biomédica*, 280-301. México: Elsevier.
72. Schunk, D. (2012,a). Conductismo. En Schunk.D., *Teorías del aprendizaje*, 71-116). México: Pearson.
73. Schunk, D. (2012,b). Constructivismo. En D. Schunk, *Teorías del aprendizaje*, 228-277. México: Pearson.
74. Styers, M., Van Zandt, P., & Hayden, K. (2018). Active Learning in Flipped Life Science Courses Promotes Development of Critical Thinking Skills. *CBE—Life Sciences Education*, 1-19.
75. ten Cate, O. (2018). Assessment of Clinical Reasoning Using the CBCR Test. En ten Cate, Custers, & Durning, *Principles and Practice of Case- based Clinical Reasoning Education: A method for preclinical students*, 85-94. Holanda: Springer Open.

76. ten Cate, O., & Loon, M. (2018). Writing CBCR cases. En ten Cate, Custers & Durning, *Principles and practice of case-based clinical reasoning education*, 95-108. Switzerland: Springer Open.
77. Verheggen, M., Muijtens, J., Van Os, J., & Schuwirth, W. (2018). Is an Angoff standard an indication of minimal competence of examinees or of judges? *Advances in Health Sciences Education*, 203-2011.
78. Victor-Chmil, J. (2013). Critical thinking versus clinical reasoning versus clinical judgment. Differential Diagnosis. *Nurse Educator*, 34-36.
79. Viesca, C., Ponce de León, M., & Sánchez, M. (s.f.). El ejercicio actual de la Medicina. *Razonamiento clínico*. Ciudad de México.
80. Villarroel, J., & Bernal, N. (2014). Razonamiento clínico: su déficit actual y la importancia del aprendizaje de un método durante la formación de la competencia clínica del futuro médico. *Rev Cienc Med*, 29-36.
81. Woods, N., Neville, A., Levins, A., Howey, E., Oczkowski, W., & Norman, G. (2006). The Value of Basic Science in Clinical Diagnosis. *Academic Medicine*, 124-127.
82. 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.

83. Zúñiga, M., & Montero, E. (2007). Teoría G: un futuro paradigma para el análisis de pruebas psicométricas. *Actualidades en Psicología*, 117-144:

ANEXO 1. RECURSOS Y CONSIDERACIONES ÉTICAS

Recursos

Para la realización del presente proyecto se utilizarán los siguientes recursos:

- Estudiantes que acepten participar voluntariamente y cuenten con los criterios de inclusión.
- Profesores que aceptaron participar voluntariamente y cumplieron con los criterios de inclusión.
- Médicos que aceptaron participar voluntariamente y cumplieron con los criterios de inclusión.
- Consumibles de papelería: hojas blancas, tinta, engrapadora, grapas, bolígrafos, lápices, equipo de cómputo, memorias USB, proyector, señaladores, cables HDMI, cables VGA, software para análisis de datos.

Consideraciones éticas

La participación en el presente proyecto se hizo de manera voluntaria, manteniendo en el anonimato la identidad de los participantes. A los profesores se les hizo entrega de una carta de invitación, un consentimiento informado, un acuerdo de confidencialidad y un aviso de privacidad. Asimismo, serán libres de retirarse en el momento que así lo consideren necesario, aún si su participación no ha concluido.

De la misma manera, esta documentación se entregará a los estudiantes, no afectará sus calificaciones y serán libres de retirarse si así lo desean.

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

Figura 1. Carta de invitación a profesores y médicos participantes.



Ciudad de México a 18 de febrero de 2021

Estimado Doctor:

El motivo de la presente es invitarle a participar en el proyecto "Diseño de una herramienta para la evaluación del pensamiento crítico en estudiantes de la asignatura de Bioquímica y Biología Molecular de la Facultad de Medicina".

El objetivo del proyecto es contar con un instrumento de evaluación que cuente con suficientes evidencias de validez y pueda aportar información sobre la situación actual de los estudiantes de la licenciatura de Médico Cirujano.

Los resultados obtenidos de este proyecto están relacionados con una de las competencias esperadas en los estudiantes al finalizar el ciclo escolar; sin embargo, estos resultados se utilizarán únicamente con fines de investigación ya que no afectarán la calificación de los estudiantes ni se utilizarán para tomar decisiones con respecto de la aprobación de la asignatura.

Para el desarrollo de este proyecto se requiere la colaboración de expertos que participarán en la elaboración y revisión de material relacionado con la asignatura y la práctica médica. Por lo que esperamos contar con su participación y compromiso.

Juan Pablo Pardo Vázquez
Profesor Titular "C" de tiempo completo
Responsable del proyecto

Figura 2. Carátula del acuerdo de confidencialidad.

ACUERDO DE CONFIDENCIALIDAD

Ciudad de México a 11 de marzo del 2021

El presente acuerdo lo celebran por una parte **Oliva Briz Tena y Juan Pablo Pardo Vázquez, nombrados de ahora en adelante, PARTE TITULAR**, personas físicas; por la otra, _____, nombrado de ahora en adelante **PARTE PARTICIPANTE**, persona física.

Las cuales en lo sucesivo serán referidas de manera conjunta como "**las partes**" y de manera individual como "**la parte**", quienes se reconocen expresa y recíprocamente con capacidad plena para obligarse, y establecen lo siguiente:

DEFINICIONES

INFORMACIÓN CONFIDENCIAL. toda información ya sea oral, impresa, o contenida en cualquier tipo de documento, archivo, o medio que pueda almacenar información independientemente de su naturaleza, relacionada con el proyecto "**Diseño de una herramienta para la evaluación del pensamiento crítico en estudiantes de la asignatura de Bioquímica y Biología Molecular de la Facultad de Medicina**".

PARTE TITULAR. Tendrán el carácter de "titular", **Oliva Briz Tena y Juan Pablo Pardo Vázquez** y toda la información confidencial relacionada con el proyecto mencionado anteriormente, será de su propiedad, aunque sea elaborada por la parte participante.

PARTE PARTICIPANTE. Tendrá el carácter de "participante" toda persona distinta de la parte titular y se obliga a guardar la información y no hacerla del conocimiento de un tercero, bajo pena de indemnizar a la parte titular en los términos y condiciones que se establecen en este convenio.

Con base en las declaraciones y definiciones anteriores, las partes se someten al tenor de las siguientes:

CLÁUSULAS

PRIMERA. OBJETO DEL ACUERDO

El presente acuerdo se refiere a la información confidencial que la **PARTE TITULAR** compartirá con la **PARTE PARTICIPANTE**, a fin de colaborar en el desarrollo del proyecto

"Diseño de una herramienta para la evaluación del pensamiento crítico en estudiantes de la asignatura de Bioquímica y Biología Molecular de la Facultad de Medicina"

Figura 3. Carátula del consentimiento informado.

CONSENTIMIENTO INFORMADO PARA LA PARTICIPACIÓN EN EL PROYECTO "DISEÑO DE UNA HERRAMIENTA PARA LA EVALUACIÓN DEL RAZONAMIENTO CLÍNICO EN ESTUDIANTES DE LA ASIGNATURA DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR EN LA FACULTAD DE MEDICINA"

Yo, _____, profesor de la asignatura de Bioquímica y Biología Molecular, en forma voluntaria, acepto participar como experto en el proyecto "Diseño de una herramienta para la evaluación del pensamiento crítico en estudiantes de la asignatura de Bioquímica y Biología Molecular en la Facultad de Medicina".

Luego de comprender en su totalidad, la información sobre dicho proyecto, entiendo que:

Mi participación consiste en la revisión de material para el desarrollo de este proyecto y toda la información será propiedad de los responsables del proyecto mencionado anteriormente.

A partir de la recepción del material tendré dos semanas para entregar la información requerida y me comprometo a cumplirlos.

La finalidad de este proyecto es elaborar una herramienta para evaluar el desarrollo de una de las competencias que se espera que los estudiantes adquieran al finalizar el ciclo escolar y que los resultados obtenidos se utilizarán únicamente con fines de investigación, por lo que no tendrán repercusiones en la calificación de los estudiantes ni se utilizarán para tomar decisiones con respecto a la aprobación de la asignatura.

No habrá ninguna sanción para mí en caso de no aceptar esta invitación.

Puedo retirarme del proyecto si lo considero conveniente a mis intereses, aún cuando el investigador responsable no lo solicite, informando mis razones para tal decisión en la Carta de Revocación respectiva, sin derecho a recuperar toda la información obtenida de mi participación ya que su utilización con fines diferentes a este proyecto puede poner en riesgo las evidencias de validez del estudio.

Violar los acuerdos de confidencialidad, pondrá en riesgo las evidencias de validez del presente proyecto y me comprometo a no hacerlo.

No haré gastos ni recibiré remuneración alguna por la participación en el estudio.

Se guardará estricta confidencialidad sobre los datos obtenidos producto de mi participación, con un número clave que ocultará mi identidad.

Puedo solicitar, en el transcurso del estudio información actualizada sobre el mismo, al Oliva Britz Tena.

También tengo acceso a las Comisiones de Investigación y Ética de la Facultad de Medicina en caso de que tenga una duda sobre mis derechos como participante en el estudio, al teléfono 56-23-22-98.

Ciudad de México a: _____

Nombre y firma del participante

TESTIGO 1

TESTIGO 2

Nombre: _____

Nombre: _____

Fecha: _____

Fecha: _____

Figura 4. Carátula del aviso de privacidad.

AVISO DE PRIVACIDAD

De conformidad con lo establecido en la Ley Federal de Protección de Datos Personales y Reglamento de Transparencia y Acceso a la Información Pública de la UNAM en posesión de los particulares, **Oliva Briz Tena y Juan Pablo Pardo Vázquez, nombrados de ahora en adelante RESPONSABLES DEL PROYECTO**, ponen a su disposición el siguiente aviso de privacidad.

Los RESPONSABLE DEL PROYECTO estarán a cargo del uso y protección de sus datos personales, en este sentido y atendiendo las obligaciones legales establecidas en la Ley Federal de Protección de Datos Personales en posesión de los particulares, a través de este instrumento se informa a los titulares de los datos, que la información que de ellos se recaba se utilizará únicamente con fines de investigación educativa.

Para llevar a cabo la finalidad descrita en el presente aviso de privacidad, se utilizarán los siguientes datos personales:

- Nombre completo
- Profesión
- Especialidad
- Años de experiencia profesional
- Años de experiencia docente
- Asignatura que imparte
- Información obtenida como resultado de su participación en el proyecto

Por otra parte, se le informa que sus datos personales no serán compartidos con ninguna autoridad, empresa, organización o persona distinta a los **RESPONSABLES DEL PROYECTO** y serán utilizados exclusivamente para los fines señalados. En caso de que la información elaborada con fines de investigación educativa se publicara, sus datos personales permanecerán en el anonimato.

Usted tiene en todo momento el derecho a conocer qué datos personales se tienen de usted, para qué se utilizan y las condiciones de uso que se les da (Acceso). Asimismo, es su derecho solicitar la corrección de su información personal en caso de que esté desactualizada, sea inexacta o incompleta (Rectificación); de igual manera, tiene derecho a que su información se elimine de este registro o bases de datos cuando considere que la misma no está siendo utilizada adecuadamente (Cancelación); así como también a oponerse al uso de sus datos personales para fines específicos (Oposición). Estos derechos se conocen como derechos ARCO.

Para el ejercicio de cualquiera de los derechos ARCO, se deberá presentar la solicitud respectiva por escrito, mediante el envío de una carta o solicitud en formato libre a la dirección de correo electrónico:

olibt.unam@gmail.com

Lo anterior también servirá para conocer el procedimiento y requisitos para el ejercicio de los derechos ARCO.

Cabe mencionar, que en cualquier momento usted puede revocar su consentimiento para el uso de sus datos personales.

La respuesta a las peticiones se dará a conocer en un plazo de 20 días hábiles.

ANEXO 2. PLAN GENERAL

PLAN GENERAL PARA EL EXAMEN DE CONOCIMIENTO GENERALES		
CONSTRUCTO	Conocimiento de las bases bioquímicas relacionadas con la diabetes tipo 2.	
Dimensión	Dominio	Constructo
Cognición	Conocimiento	Identificar o reconocer el constructo y reconocer las ideas principales.
	Comprensión	Entender la información y su significado, trasladarla a nuevos contextos y solucionar problemas utilizando conocimiento previo.
	Aplicación	Utilizar el conocimiento aprendido en situaciones nuevas.

CARACTERÍSTICAS DEL INSTRUMENTO

Propósito del examen	Evaluar conocimiento factual y conceptual, así como su aplicación.
Población objetivo	Estudiantes de la asignatura de Bioquímica y Biología Molecular que hayan concluido el tercer bloque de la licenciatura de médico cirujano.
Alcances de los resultados de la evaluación	<p>Inferir si los estudiantes conocen, comprenden y aplican el conocimiento de las bases bioquímicas relacionadas con la diabetes tipo 2.</p> <p>Inferir si los estudiantes relacionan las bases bioquímicas correspondientes a la diabetes tipo 2 con esta enfermedad.</p> <p>A partir de estos resultados se podrá fortalecer la formación de los estudiantes y proponer estrategias específicas para reforzar estos conocimientos.</p>
Limitaciones de los resultados de la evaluación	<p>A partir de los resultados de este examen no se puede inferir:</p> <p>Si los estudiantes tienen los conocimientos suficientes para estar en segundo año.</p> <p>Si los estudiantes son aptos para ser médicos.</p> <p>Si los estudiantes son inteligentes.</p> <p>Cuál es el mejor grupo del primer año de la licenciatura.</p>
Cobertura del examen	Institucional

Tipo de Evaluación	Diagnóstica
Impacto del examen	Bajo
Referente de calificación	Criterial
Sensibilidad a la instrucción formal	Alta
Responsable de la aplicación	OBT.MG
Responsable de la convocatoria	OBT.MG
Periodos de aplicación al año	Aplicación única, al término del tercer bloque.
Descripción de las condiciones de la aplicación	<p>Se aplicará en línea, a todos los estudiantes en el mismo horario; cada alumno accederá al examen desde su ordenador o dispositivo, usando su propio servicio de internet. El examen tendrá una duración de 1 hora, será supervisado por los profesores titulares o adjuntos del grupo en una sesión de zoom con cámaras encendidas.</p> <p>Se especificará que es un examen diagnóstico y los alumnos tendrán aviso previo del mismo.</p>

Número y duración de las sesiones que se requieren para la resolución del examen	Una sesión, 1 hora
Características de la sesión de aplicación	Medianamente controlada. Misma hora para todos, inicia y termina al mismo tiempo para todos, sesión en zoom, cámaras encendidas.
Observaciones para la aplicación	Se iniciará la aplicación con una explicación sobre el objetivo de la prueba, alcance y consecuencias de la misma.
Taxonomía	Bloom
Dominios del constructo	<p>Conocimiento: identificar las bases bioquímicas de la diabetes tipo 2.</p> <p>Comprensión: reconocer los puntos clave de las bases bioquímicas de la diabetes tipo 2.</p> <p>Aplicación: solucionar problemas con base en el conocimiento previo.</p>
Número de instrumentos	Uno, una sola versión.
Formato de ítems	Opción múltiple

Número de opciones de respuesta	de	3 y sólo una es correcta
Longitud del examen	del	34 preguntas
CALIFICACIÓN		
Escala en que se reporta la calificación	de	Número de aciertos.
Tipo de calificación	de	Calificación global
Método para el establecimiento de los puntos de corte	de	Angoff
Reporte por niveles de desempeño	por	Se entregará a los sustentantes un reporte del número total de aciertos.
Modelo de calificación	de	Por puntaje de aciertos: Reactivo contestado correctamente = 1 punto Reactivo contestado erróneamente = 0 puntos

	No se penalizan las respuestas erróneas o preguntas sin responder.
Descripción de la base de datos o plantilla de calificación	Formato Excel con identificador del sustentante y número de aciertos de acuerdo con la estructura del instrumento.
Reportes de resultados	Individual e institucional Se elaborará un Reporte técnico de los resultados de la aplicación que servirá de insumo a los responsables del instrumento para el Mantenimiento del instrumento.
Responsable de la elaboración de los reportes de resultados	OBT.MG

		razones que sustenten su punto de vista. Identificar y diferenciar premisas que permitan sustentar una conclusión tras el análisis de una expresión.
CARACTERÍSTICAS DEL INSTRUMENTO		
Propósito del examen	Evaluar el pensamiento crítico en estudiantes de la asignatura de Bioquímica y Biología Molecular.	
Población objetivo	Estudiantes de la asignatura de Bioquímica y Biología Molecular que hayan concluido el tercer bloque de la licenciatura de médico cirujano.	
Alcances de los resultados de la evaluación	<p>Inferir si los estudiantes relacionan las bases bioquímicas de la diabetes tipo 2 con esta enfermedad.</p> <p>A partir de estos resultados se podrá fortalecer la formación de los estudiantes y/o diseñar estrategias específicas para reforzar estos conocimientos.</p>	
Limitaciones de los resultados de la evaluación	<p>A partir de los resultados de este examen no se puede inferir:</p> <p>Si los estudiantes tienen los conocimientos suficientes para estar en segundo año.</p>	

	<p>Si los estudiantes son aptos para ser médicos.</p> <p>Si los estudiantes son inteligentes.</p> <p>Cuál es el mejor grupo del primer año de la licenciatura.</p>
Cobertura del examen	Institucional
Tipo de Evaluación	Diagnóstica
Impacto del examen	Bajo
Referente de calificación	Criterial
Sensibilidad a la instrucción formal	Alta
Responsable de la aplicación	OBT.MG
Responsable de la convocatoria	OBT.MG
Periodos de aplicación al año	Aplicación única, al término del tercer bloque.
Descripción de las condiciones de la aplicación	<p>Se aplicará en línea, a todos los estudiantes en el mismo horario, cada alumno accederá al examen desde su ordenador o dispositivo, usando su propio servicio de internet, el examen tendrá una duración de 1 hora, será supervisado por los profesores titulares o adjuntos del grupo en una sesión de zoom con cámaras encendidas.</p> <p>Se especificará que es un examen diagnóstico y los alumnos tendrán aviso previo del mismo.</p>

Número y duración de las sesiones que se requieren para la resolución del examen	Una sesión, 1 hora
Características de la sesión de aplicación	Medianamente controlada. Misma hora para todos, inicia y termina al mismo tiempo para todos, sesión en zoom, cámaras encendidas.
Observaciones para la aplicación	Se iniciará la aplicación con una explicación sobre el objetivo de la prueba, alcance y consecuencias de la misma.
Taxonomía	Bloom con base en los dominios de interpretación y análisis de la definición de pensamiento crítico de Facione.
Dominios del constructo	Análisis: Divide el conocimiento en partes y muestra relaciones entre ellas. Aplicación: utiliza el conocimiento aprendido en situaciones nuevas.
Número de instrumentos	Uno, una sola versión.
Formato de ítems	Opción múltiple
Número de opciones de respuesta	5 columnas con una lista de opciones cada una.
Longitud del examen	7 preguntas.

CALIFICACIÓN	
Escala en que se reporta la calificación	Número de aciertos y escala del 1-10
Tipo de calificación	Calificación Global
Puntos de corte	Opinión de expertos.
Método para el establecimiento de los puntos de corte	Angoff
Reporte por niveles de desempeño	Se entregará a los sustentantes un reporte del número de aciertos global.
Modelo de calificación	<p>Por puntaje de aciertos:</p> <p>Reactivo contestado correctamente = 1 punto</p> <p>Reactivo contestado erróneamente = 0 puntos</p> <p>No se penalizan las respuestas erróneas o preguntas sin responder.</p>
Descripción de la base de datos o plantilla de calificación	Formato Excel con identificador del sustentante y número de aciertos de acuerdo con la estructura del instrumento.
Reportes de resultados	Individual e institucional

	Se elaborará un reporte técnico de los resultados de la aplicación que servirá de insumo a los responsables del instrumento para el mantenimiento del instrumento.
Responsable de la elaboración de los reportes de resultados	OBT.MG

ANEXO 3. CALIDAD DE LOS REACTIVOS

Evaluación de la calidad de los reactivos 1 – 15 del examen de conocimientos generales.

Dominio a evaluar	Juez	Número de reactivo															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1.- ¿La cantidad de texto en el tallo es adecuada para su comprensión?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	N O	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI	SI	SI	SI
2.- ¿La pregunta o instrucción se encuentra redactada con claridad?	2	S	SI								N O	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	N O	SI	SI	SI	SI	SI	SI	N O	SI	SI	N O	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
3.- ¿El reactivo cuenta con una gramática, puntuación y ortografía correctas?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	N O	N O	N O	SI	N O	SI	SI	SI	SI	N O	SI	SI	N O	SI	SI	SI
	6	N O	SI	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
4.- ¿El tallo del reactivo plantea la idea central?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
5.- ¿El reactivo presenta un solo resultado de aprendizaje?	2	N O	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	N O	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	

	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
6.- ¿El reactivo presenta un solo contenido temático?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
		6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
7.- ¿La semántica utilizada está de acuerdo con el contenido del programa académico?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
		6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
8.- ¿El reactivo cuenta únicamente con una respuesta?	2	SI	N O								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
		6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
9.- ¿Las opciones son independientes entre sí?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
		6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
10.- ¿El contenido evaluado está en relación con la especificación del reactivo?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
		6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
11.- ¿Las opciones son similares en cuanto a estructura gramatical, contenido y extensión?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	N O	SI	SI	SI	SI	SI	SI	N O	SI	SI	N O	SI	N O	NO
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
		6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
12.- ¿Las opciones evitan dar pistas sobre la respuesta correcta?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	N O	NO	
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI	SI	N O	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI	SI	SI
		2	SI	SI								SI	SI	SI	SI	SI	SI

13.- ¿Los distractores son plausibles, es decir, no se descartan por inferencia lógica o sentido común?	3	SI	NO	SI	SI	SI											
	4	SI															
	5	SI	SI	NO	SI	SI	SI	SI	SI	NO	NO	SI	NO	SI	NO	NO	NO
	6	SI															
14.- ¿El reactivo cuenta con tres o cuatro opciones de respuesta?	2	SI	SI								SI	SI	SI	SI	SI	SI	
	3	SI															
	4	SI															
	5	SI	NO	SI	SI	SI	SI	SI									
	6	SI															

**Los espacios vacíos corresponden a la opinión de los jueces que fueron autores del reactivo por lo que no fue tomada en cuenta.

Evaluación de la calidad de los reactivos 16 – 30 del examen de conocimientos generales.

Dominio a evaluar	Juez	Número de reactivo														
		1	1	1	1	2	2	2	2	2	2	2	2	2	2	3
		6	7	8	9	0	1	2	3	4	5	6	7	8	9	0
1.- ¿La cantidad de texto en el tallo es adecuada para su comprensión?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	NO	SI										
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
2.- ¿La pregunta o instrucción se encuentra redactada con claridad?	2	SI	SI	SI	SI	SI	SI	NO	SI							
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	NO	SI								
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
3.- ¿El reactivo cuenta con una gramática, puntuación y ortografía correctas?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	NO	SI	SI
	5	SI	SI	SI	SI	SI	NO	SI								
	6	SI	SI	SI	SI	NO	SI									
4.- ¿El tallo del reactivo plantea la idea central?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI

	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
5.- ¿El reactivo presenta un solo resultado de aprendizaje?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
6.- ¿El reactivo presenta un solo contenido temático?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
7.- ¿La semántica utilizada está de acuerdo con el contenido del programa académico?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
8.- ¿El reactivo cuenta únicamente con una respuesta?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
9.- ¿Las opciones son independientes entre sí?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
10.- ¿El contenido evaluado está en relación con la especificación del reactivo?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI
11.- ¿Las opciones son similares en cuanto a estructura gramatical, contenido y extensión?	2	SI	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	N O	SI						
	6	SI	SI	N O	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
12.- ¿Las opciones evitan dar pistas sobre la respuesta correcta?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI

	6	SI	SI	N O	N O	SI										
13.- ¿Los distractores son plausibles, es decir, no se descartan por inferencia lógica o sentido común?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
14.- ¿El reactivo cuenta con tres o cuatro opciones de respuesta?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	3	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	5	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI

Evaluación de la calidad de los reactivos 16 – 30 del examen de conocimientos generales.

Dominio a evaluar	Juez	Número de reactivo															
		31	32	33	34	35	36	37	38	39	40	41	42	43	44		
1.- ¿La cantidad de texto en el tallo es adecuada para su comprensión?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	3	SI				SI											
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI	SI	
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI					
2.- ¿La pregunta o instrucción se encuentra redactada con claridad?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	3	SI				SI											
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI	SI	
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI					
3.- ¿El reactivo cuenta con una gramática, puntuación y ortografía correctas?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	3	SI				SI											
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI	SI	
	6	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI					
4.- ¿El tallo del reactivo plantea la idea central?	2	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	3	SI				SI											
	4	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI	SI	

	6	SI													
5.- ¿El reactivo presenta un solo resultado de aprendizaje?	2	SI													
	3	SI				SI									
	4	SI													
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI
	6	SI													
6.- ¿El reactivo presenta un solo contenido temático?	2	SI													
	3	SI				SI									
	4	SI													
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI
	6	SI													
7.- ¿La semántica utilizada está de acuerdo con el contenido del programa académico?	2	SI													
	3	SI				SI									
	4	SI													
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI
	6	SI													
8.- ¿El reactivo cuenta únicamente con una respuesta?	2	SI													
	3	SI				SI									
	4	SI													
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI
	6	SI	SI	SI	SI	N	SI	SI	SI	SI	SI				
					O										
9.- ¿Las opciones son independientes entre sí?	2	SI													
	3	SI				SI									
	4	SI													
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI

	6	SI													
10.- ¿El contenido evaluado está en relación con la especificación del reactivo?	2	SI													
	3	SI				SI									
	4	SI													
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI
	6	SI													
	11.- ¿Las opciones son similares en cuanto a estructura gramatical, contenido y extensión?	2	SI												
3		SI				SI									
4		SI													
5		SI	SI	SI	SI			SI			N	SI	SI	SI	SI
6		SI	SI	SI	N	SI	SI	SI	SI	N	SI				
					O					O					
12.- ¿Las opciones evitan dar pistas sobre la respuesta correcta?	2	SI													
	3	SI				SI									
	4	SI													
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI
	6	SI	SI	SI	N	SI	SI	SI	SI	SI	SI				
					O										
13.- ¿Los distractores son plausibles, es decir, no se descartan por inferencia lógica o sentido común?	2	SI													
	3	SI				SI									
	4	SI	I	SI											
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI
	6	SI													
	2	SI													
	3	SI				SI									

14.- ¿El reactivo cuenta con tres o cuatro opciones de respuesta?	4	SI													
	5	SI	SI	SI	SI			SI			SI	SI	SI	SI	SI
	6	SI													

ANEXO 4. PSICOMETRÍA

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

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