

Validación del constructo de un banco de reactivos para evaluar el egreso en la licenciatura en Odontología

Juárez Méndez, Rocío; Castañeda Figueiras, Sandra

Validación del constructo de un banco de reactivos para evaluar el egreso en la licenciatura en Odontología

Psicología Iberoamericana, vol. 26, núm. 1, 2018

Universidad Iberoamericana, Ciudad de México, México

Disponible en: <https://www.redalyc.org/articulo.oa?id=133959553007>

Validación del constructo de un banco de reactivos para evaluar el egreso en la licenciatura en Odontología

Construct validation of items bank to assess the graduation in dentistry

Rocío Juárez Méndez rociojm2012@gmail.com

Universidad Nacional Autónoma De México, México

Sandra Castañeda Figueiras

Universidad Nacional Autónoma De México, México

Resumen: Nuestro propósito fue validar constructos subyacentes al examen que, en el ámbito nacional, evalúa conocimientos y habilidades de egresados de la licenciatura en Odontología. Para contar con sustento empírico sobre los resultados en el examen, el estudio implicó: a) seleccionar 24 reactivos procedentes de un banco mayor (360) que cubrieron los requisitos de calibración logística de dos parámetros, Zimowski, Muraki, Mislevy y Bock (2003), representativos de las tres áreas del examen. La muestra fue de 535 egresados; b) se caracterizaron las fuentes de contenido de los reactivos por nivel de complejidad: demanda cognitiva y tipos de conocimiento subyacentes; c) se validó el constructo primero mediante un análisis factorial exploratorio y posteriormente mediante un análisis factorial confirmatorio, con esto se puso a prueba un modelo multirrasgo-multimétodo. El modelo utilizó 24 reactivos calibrados y fueron asociado a una doble condición: demanda cognitiva y tipo de conocimiento. Los índices de ajuste práctico mostraron que no hay diferencias significativas entre modelos: $cfi = .96$, $rmsea = .03$, intervalo de confianza $rmsea = (.020, -.035)$, $cfi = .95$ y $agfi = .94$. La estructura factorial del modelo muestra validez convergente de constructo la demanda cognitiva aplicar y el tipo de conocimiento procedimental $r = .52$. La validez divergente entre dimensiones mostró $r = -.07$ entre los constructos comprender y resolver problemas, $r = .02$ entre los tipos de conocimiento conceptual y procedimental, $r = .13$ entre los constructos comprender y aplicar, y $r = -1.00$ entre las demandas cognitivas aplicar y resolver problemas, así es posible establecer inferencias entre lo que se comprende y se resuelve entre lo procedimental y lo conceptual.

Palabras clave: validez de constructo, demanda cognitiva, fuentes de contenido, modelo multirrasgo, multi-método.

Abstract: Our purpose was to validate the constructs underlying test that, nationally, evaluates knowledge and skills of graduates from the Bachelor of Dentistry. To have empirical evidence about the test results, the study involved: a) Select 24 items from a scale test (360 items) covering logistics requirements two calibration parameters, Zimowski, Muraki, Mislevy and Bock (2003), representative of the three areas of the test. The sample was 535 graduates; b) sources of content reagents were characterized by level of complexity: cognitive demand and types of underlying knowledge; finally, c) the construct was validated using exploratory factor analysis and confirmatory factor analysis following a multitrait-multimethod model. The model used 24 reagents were calibrated and linked to two conditions, cognitive demand and type of knowledge. The practical adjustment indexes showed no significant differences between models: $cfi = .96$, $rmsea = .03$, $rmsea$ confidence interval = $(.020, -.035)$, $cfi = .95$ and $agfi = .94$. The factorial structure model shows convergent construct validity cognitive demand and implement the type of procedural knowledge $r = .52$. The divergent validity between dimensions showed $r = -.07$ between constructs understand and solve problems $r = .02$ between the types of conceptual and procedural knowledge, $r = .13$ between constructs understand and apply, $r = -1.00$ between the cognitive demands implement and solve problems, and

Psicología Iberoamericana, vol. 26, núm. 1, 2018

Universidad Iberoamericana, Ciudad de México, México

Recepción: 17 Agosto 2017
Aprobación: 12 Abril 2018

Redalyc: <https://www.redalyc.org/articulo.oa?id=133959553007>

can make inferences from what is understood and resolved between procedural and conceptual.

Keywords: construct validity, cognitive demand, content sources, multitrait-multi, method model.

INTRODUCCIÓN

El entorno laboral exige profesionistas mejor preparados para responder a las exigencias del mundo actual. En este sentido, la evaluación del aprendizaje tiene un papel preponderante, porque la comunidad educativa y la sociedad en su conjunto tienen altas expectativas respecto a los resultados que de ella se derivan, en virtud de que los instrumentos de medición deben ser una herramienta para legitimar los conocimientos y habilidades de quienes han concluido un programa educativo.

Sin embargo, a juicio de los expertos, existen deficiencias en los procesos de evaluación educativa, como la falta de congruencia entre los objetivos y las metas de la evaluación y los propios instrumentos; es decir, con frecuencia los instrumentos no tienen relación con los propósitos para los cuales fueron diseñados, por lo que repercuten en la imposibilidad de utilizar los resultados de la evaluación para el análisis y elaboración de conclusiones acerca de los conocimientos, habilidades y competencias de quienes concluyen un nivel educativo, o bien identificar con precisión quiénes son capaces de resolver problemas novedosos y complejos. Esta situación ha sido discutido desde hace varias décadas, en los años cuarenta Tyler (1969) planteó la necesidad de una evaluación científica que fuera de utilidad para perfeccionar la calidad de la educación, pues consideraba que los exámenes tenían baja calidad y las calificaciones eran imprecisas, debido a la heterogeneidad de los criterios empleados y al divorcio entre lo que se enseña y las metas de la instrucción. Asimismo Cronbach y Snow (1969) criticaron el uso de las evaluaciones por su carencia de relevancia y utilidad, afirmando que la evaluación debía basarse en objetivos explícitos que ayudaran a los docentes a tomar decisiones sobre cómo educar.

Para Whitely (1983) y Embretson (1985), la información de procesamiento de la ejecución de las tareas que enfrentaban los sujetos evaluados implicaba cambios no sólo en las preguntas que se les pedía, sino también en el tipo de datos que se consideraban pertinentes. ¿Pero cómo se logra establecer este vínculo? En este sentido, la validez de constructo es una poderosa herramienta que centra su atención en una función indispensable de la teoría psicológica, en la elaboración de las pruebas y en la necesidad de formular hipótesis que aporten elementos para aceptarlas o rechazarlas durante el proceso de validación.

La validación de constructo requiere de la acumulación gradual de diversas fuentes de información; cada constructo se deriva de las interrelaciones establecidas entre las medidas conductuales y se forma para organizar y dar cuenta de las concordancias observadas en las respuestas. Cualquier información que arroje luz sobre la naturaleza de los datos considerados y sobre las condiciones que afectan su desarrollo

y manifestaciones constituye una evidencia apropiada para su validación (Anastasi & Urbina, 1998).

Por esta razón, ya no se admite fraccionar la validez en tres tipos. De acuerdo con Messick (1995), el concepto tradicional de validez que la dividía en validez de contenido, criterio y constructo es fragmentado e incompleto, principalmente porque no se consideran las evidencias de los puntajes y lo que implica la interpretación de su significado. El concepto vigente y unificado de validez interrelaciona estos temas como aspectos fundamentales de una teoría más completa de la validez de constructo que aborda tanto el significado como la puntuación y los valores sociales en la interpretación de la prueba y el uso de la prueba. Es decir, la validez unificada integra las consideraciones de contenido, criterios y las consecuencias en un marco de construcción de la comprobación empírica de las hipótesis racionales sobre el significado de la puntuación y las relaciones teóricamente relevantes, incluyendo una aplicación de carácter científica. El autor distingue seis puntos de la validez de constructo como un medio para abordar aspectos centrales que llevan implícita la noción de validez como un concepto integral y funcionan como criterios de validez general o estándares para todos los centros educativos y psicológicos de medición, incluyendo evaluación de las ejecuciones, que se examinan a detalle debido a su creciente énfasis en la educación y los ajustes que deben ser empleados. Estos aspectos son el contenido, los modelos teóricos y de procesos, la estructura, la generabilidad, los aspectos externos y los aspectos consecuentes de la validez de constructo.

El primer punto se refiere a la pertinencia y representatividad del contenido, así como a la calidad técnica de los reactivos.

Los modelos teóricos y de procesos se refieren a los fundamentos teóricos para observar la consistencia en las respuestas de la prueba; incluidos los modelos de los procesos de ejecución de la tarea junto con pruebas empíricas, para lo cual se sugiere realizar observaciones de campo sobre la ejecución de los examinados, para registrar de manera directa las estrategias de solución y el tiempo requerido para realizar la tarea. La estructura establece la relación entre la puntuación con las manifestaciones del comportamiento, es decir, se analiza la fidelidad de la ejecución de la tarea y cómo se combina para producir una puntuación.

La generalibilidad examina la medida con las que las puntuaciones se pueden extrapolar hacia otros grupos de población y cómo las comparaciones pueden predecir la interpretación de los resultados.

Los aspectos externos se refieren a validez convergente y divergente de las pruebas, hasta qué punto los resultados de las mediciones convergen con otras medidas y constructos, y hasta qué punto divergen de las mismas.

Los aspectos consecuentes de la validez de constructo evalúan el valor de la puntuación, las implicaciones de interpretación como base para la acción, así como las consecuencias reales y potenciales del uso de la prueba. En este tenor, Embretson (1999) enfatiza que los elementos derivados de la investigación cognitiva para implantar los nuevos objetivos de aprendizaje son necesarios para generar modelos, procedimientos, técnicas e instrumentos cognitivos y psicométricos que

permitan diseñar mediciones que aseguren la mayor validez posible y eviten sesgos derivados de aproximaciones psicométricas inadecuadas. Sin embargo, este cambio va más allá de un simple reemplazo del método, sus implicaciones representan una nueva concepción del aprendizaje académico, en el cuál la comprensión, el pensamiento, la solución de problemas y el conocimiento reflexivo son interdependientes del contenido por aprender y del contexto del aprendizaje mismo.

Por lo anterior, para Embretson (1999) se requiere una teoría de construcción de exámenes que tome en cuenta los siguientes aspectos, estrechamente relacionados:

El diseño de la observación, para construir y organizar materiales, tareas y reactivos que permitan revelar los mecanismos que los examinados usan para responder, así como para generar evidencias suficientes que permitan inferir en los procesos, las estructuras de conocimientos y las estrategias que poseen o no los examinados.

El diseño de la medición, para definir el objeto de medida y las unidades de análisis para describir el procedimiento o grupo de procedimientos utilizados para asignar un valor o una categoría a un objeto de medida y para establecer las formas de asegurar la precisión del procedimiento para asignar un valor o categoría.

Tomando en cuenta este cambio de paradigma, una de las investigaciones más importantes realizadas bajo estos supuestos fue dirigida por Castañeda en 2004, quien, junto con un grupo de académicos del Laboratorio de Evaluación y Fomento del Desarrollo Cognitivo y el Aprendizaje Complejo del Posgrado de la Facultad de Psicología de la unam e investigadores de la Universidad de Sonora, se interesó por validar los constructos que hipotéticamente subyacen a fuentes de contenido de ítems objetivos de tres exámenes de egreso de la licenciatura, con el propósito de entender el proceso de responder a este tipo de examinación.

Este grupo de investigadores planteó que cualquier examen debía asegurar que tanto los constructos subyacentes, como la evidencia recabada reflejaran válidamente lo que se desea medir, de otra manera existe el riesgo de que la interpretación de los resultados contenga explicaciones perniciosas para el sustentante, autoridades educativas y todo aquel interesado en el tema.

Los investigadores establecieron como objetivo principal validar los puntajes que hipotéticamente subyacen a las fuentes de contenido de los bancos de reactivos de los exámenes de egreso de las licenciaturas de Contaduría, Pedagogía y Psicología.

Las fuentes de contenido analizadas fueron las operaciones cognitivas requeridas para responder el reactivo y el campo de conocimiento en el que se evalúa la información que el reactivo presenta.

Para el caso de la validación de constructos de la licenciatura en Psicología un segundo objetivo fue avanzar en la validez de constructo previamente, lo cual significa generar una red de evidencias múltiples, entre otras, las de contenido y las de estructura interna de la examinación.

Para tal efecto, los investigadores optaron por hacer uso de la aproximación multirasgo-multimétodo, por la capacidad que este análisis

tiene para medir dos o más rasgos o características. Los rasgos pueden ser habilidades, actitudes, operaciones cognitivas, conductas o características de personalidad; en tanto los métodos se refieren a medidas o situaciones variadas, tales como el campo de conocimiento a evaluar, los formatos de reactivos utilizados, el tipo de conocimiento, entre otros.

Los resultados de esta investigación concluyeron que la estrategia utilizada permitió recoger varianza que explicó el proceso de responder en función de dos fuentes de contenido incluidas en los ítems utilizados. También permitió modelar las relaciones estructurales entre los constructos hipotetizados, de manera integral y con carácter confirmatorio. Así se hizo factible entender cómo la interacción entre las fuentes de contenido investigadas explica o no los resultados del proceso de responder en un banco intencional de examinación a gran escala en una disciplina en particular.

Los modelos obtenidos, además de verificar la validez de constructo de las medidas analizadas, permitieron corroborar la pertinencia de estudiar el proceso de responder bajo la perspectiva del análisis factorial confirmatorio con matriz de operación cognitiva-multicampo de conocimiento.

En el campo de la neuropsicología Delis, Jacobson, Bondi, Hamilton y Salmón (2003) utilizaron correlación de Pearson y análisis factorial para evaluar la validez de constructo de la escala de aprendizaje verbal de California (cvlt). El objetivo principal de esta investigación fue enfatizar sobre la importancia de la aplicación apropiada del análisis factorial para el proceso de validez de constructo. La discusión de los resultados se enfoca en el método de la varianza múltiple y cómo esto puede distorsionar los resultados del análisis factorial.

Los autores afirman que en los últimos años ha habido un aumento significativo en el uso del análisis factorial confirmatorio, para evaluar la validez de constructo de las pruebas neuropsicológicas. En el análisis factorial exploratorio se debe asumir que todos los factores comunes son correlacionados (o en algunas aplicaciones, no correlacionados), todas las variables observadas se ven directamente afectadas por todos los factores comunes, son factores únicos no correlacionados el uno con el otro, todas las variables observadas se ven afectados por un único factor, y todos los factores comunes únicos no se correlacionan con todos los factores (Long, 1983). El análisis factorial confirmatorio permite al investigador superar las limitaciones del análisis factorial exploratorio (Kline, 1998; Long, 1983). Estas limitaciones determinan qué pares de factores comunes están correlacionados, en cuáles se observan variables que se ven afectadas por factores comunes y cuáles son pares de factores únicos que están correlacionados (Long, 1983).

En resumen, el trabajo de Delis y sus colaboradores aclara la necesidad de considerar cuidadosamente tanto la población objetivo como las variables antes de realizar la selección de procedimientos correlacionales, factoriales o analíticos. Los evaluadores deben tener en cuenta los efectos potenciales de las diferentes variables sobre las correlaciones entre las puntuaciones de la prueba, y ser conscientes de los efectos de

distorsión que puede ser ocasionada por el método de varianza causada por diversas decenas de factores basados en el mismo procedimiento de ensayo o prueba de estímulos. El análisis factorial sigue siendo una herramienta notable para evaluar la validez de constructo de las pruebas neuropsicológicas, cuando se lleva a cabo adecuadamente. El análisis factorial confirmatorio puede ampliar significativamente al análisis de la validez de constructo, al permitir contrastarlo con otros modelos estadísticos. Los evaluadores pueden emplear un enfoque cognitivo experimental en grupos de sujetos homogéneos. En síntesis, los métodos de varianza, tales como el análisis factorial, no representan técnicas inadecuadas para la validez de constructo, más bien lo inadecuado es el uso inapropiado de las técnicas.

El objetivo de este trabajo es validar los constructos del examen para evaluar el egreso de la licenciatura en Odontología mediante el análisis de las fuentes de contenido, procesos cognitivos y tipos de conocimiento, así como su relación con los reactivos.

MÉTODO

Participantes

Se contó con las respuestas de 495 participantes de 54 distintas instituciones educativas de todo el país, 66% tenía un rango de edad entre 22 y 25 años al momento de egresar, de los cuales 64% eran mujeres y 36% hombres; 54% tuvo un promedio de egreso entre 8 y 8.9.

Diseño de Investigación

Se realizó un diseño de investigación no experimental que fue desarrollado en varias fases. Posterior al análisis psicométrico de los reactivos, lo siguiente fue categorizar las áreas de contenido de la prueba y las áreas cognitivas a evaluar, utilizando una tabla bidimensional denominada Escala de Valoración de Fuentes de Contenido de Reactivos Objetivos de Castañeda (2004), con el propósito de caracterizar cada ítem con base en las diferentes fuentes de contenido que en él se incluyen. El objetivo de la escala es obtener datos empíricos de la dificultad apriorística de los ítems. Adicionalmente, permite identificar los procesos cognitivos y tipos de conocimiento requeridos para responder cada reactivo.

PROCEDIMIENTO

Análisis de la fuente de contenido: Procesos cognitivos

Se estudiaron los procesos cognitivos con el propósito de obtener información acerca de los procesos de pensamiento de los sujetos para resolver tareas que van de lo simple a lo complejo. En este sentido, los procesos cognitivos considerados en esta investigación son:

comprensión, que se define como la capacidad del examinado para identificar, clasificar, ordenar o jerarquizar la información que aparece en el reactivo; aplicación, que corresponde a la capacidad de hacer uso de conceptos, principios, procedimientos, técnicas e instrumentación en tareas profesionales o rutinarias; y solución de problemas, operación cognitiva que hace referencia a la capacidad del examinado para evaluar e integrar conceptos, principios, procedimientos, métodos, técnicas, estructuras de tareas o planes de acción en la resolución de planteamientos de situaciones problemáticas y para identificar y corregir errores en soluciones previamente establecidas.

Análisis de la fuente de contenido: Tipo de conocimiento

En esta fuente de contenido los reactivos son analizados en función del tipo de conocimiento posible a evaluar respecto a una tarea dada: el conocimiento factual, que corresponde con los componentes básicos de un dominio que se aprende memorísticamente; el conocimiento conceptual, que se refiere a las interrelaciones entre elementos, la manera en que éstos funcionan y su significado esencial, características definitorias y reglas que lo componen, conceptos, principios y explicaciones; o el conocimiento procedimental, que consiste en aplicar procedimientos para reconocer patrones, como son sensorio-motrices, perceptuales o semánticos, así como secuencia de acciones, tales como la resolución de ecuaciones, elaboración de diagnósticos, practicar algún deporte, entre otras.

Obtención del modelo factorial exploratorio. Una vez determinadas las características de los reactivos mediante el análisis de fuentes de contenido, se decidió realizar un análisis factorial exploratorio para reducir el número de variables y agrupar en componentes aquellas que están altamente correlacionadas, tratando de conservar la mayor información sobre los constructos analizados. Se empleó el método de extracción de ejes principales, con el fin de identificar un número menor de variables, las cuales se configuran en componentes, con base en la varianza común.

Las dimensiones hipotéticas se muestran en la Tabla 1. Para probar las dimensiones hipotetizadas se diseñó un modelo con el programa EQS, versión 6.1. (Bentler, 2006).

Tabla 1
Contenidos del examen

Área	Subárea
Formación básica	Anatomía y fisiología
	Microbiología y bioquímica
	Genética e histoembriología
	Patología y farmacología
	Bioética
Formación instrumental	Clínica básica
	Odontología social
	Legal-administrativo
Formación operativa	Operatoria dental
	Exodoncia y cirugía
	Endodoncia
	Odontopediatría
	Prostodoncia
	Periodoncia
	Ortodoncia
	Odontogeriatría

Tabla 1
Contenidos del examen

RESULTADOS

Tratamiento de los Datos

De forma individual, los reactivos fueron asignados a los niveles de cada fuente contenido, posteriormente éstas se interreacionaron, con lo cual se obtuvieron los resultados que se muestran en la Tabla 2.

La operación cognitiva más utilizada es comprender. Se relaciona directamente con el tipo de conocimiento conceptual y destaca que el conocimiento procedimental ocupa una porción baja dentro del examen.

Identificación de la Matriz Multirrasgo-Multimétodo

Derivado del análisis factorial exploratorio se propuso un modelo teórico conformado por 24 variables observadas que se organizan en x factores, el

criterio para seleccionar las variables finales se basó en las cargas factoriales significativas que además conformaban un factor.

Las variables corresponden a las áreas de formación básica [anatomía y fisiología (ana y fis), microbiología y bioquímica (mic y boiq), genética e histoembriología (genética)], formación instrumental [clínica básica (clínica), odontología social (odon soc)], y formación operativa [exodoncia y cirugía (exodon), endodoncia (endodon), odontopediatria (odontope), prostodoncia (prostodo)].

Los índices globales indican que el ajuste del modelo es adecuado, el índice de razón de verosimilitud (χ^2) fue de 471.80 con 224 grados de libertad y significancia de .00. Es importante considerar que el modelo fue estimado con una muestra muy alta compuesta por 495 sujetos, por lo que es importante analizar el resto de los índices.

Respecto al índice de ajuste comparativo (cfi) se obtuvo un valor de .88, lo cual indica que se encuentra en un nivel aceptable, al igual que el error de la raíz cuadrada media de aproximación (rmsea) que fue igual a .05.

El modelo se considera confiable al haberse obtenido valores de .70 tanto en Alfa de Cronbach como en el coeficiente de confiabilidad (ρ), lo que indica que son aceptables para el modelo.

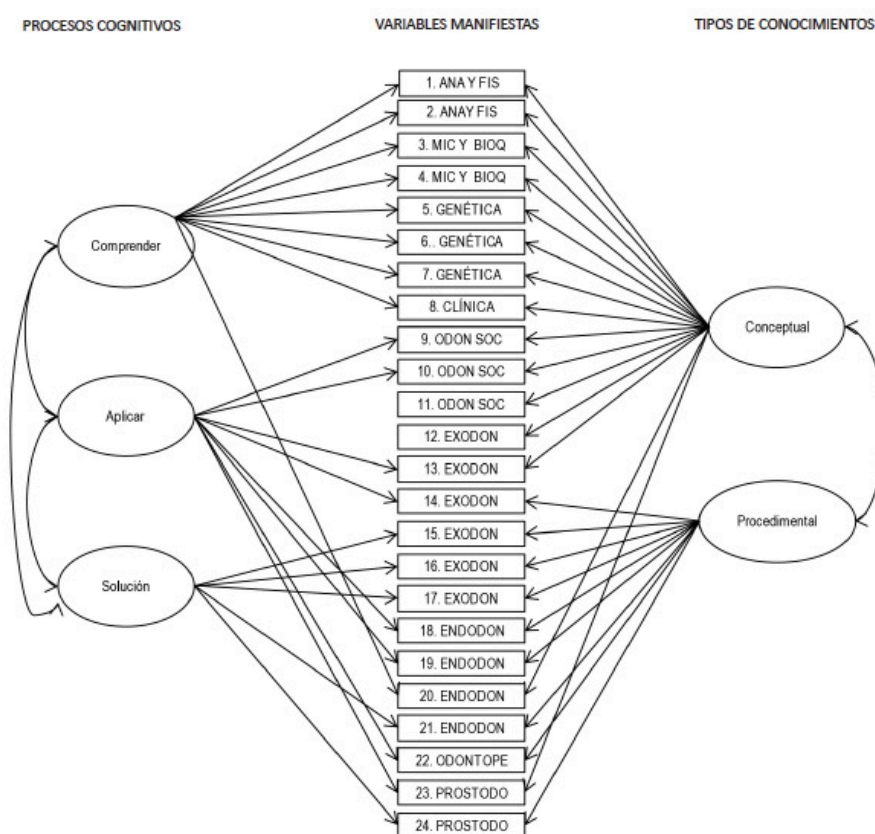


Figura 1
Modelo teórico de procesos cognitivos y tipos de conocimientos.

Figura 1
Modelo teórico de procesos cognitivos y tipos de conocimientos.

Tabla 2
Porcentaje de reactivos por operación cognitiva y tipo de conocimiento a evaluar

Área	Operación cognitiva			Tipo de conocimiento	
	Comprender	Aplicar	Resolver	Conceptual	Procedimental
Formación básica	30%	0%	0%	30%	0%
Formación instrumental	12%	10%	0%	16%	7%
Formación operativa	9%	26%	12%	20%	27%
Total	51%	36%	12%	66%	34%

Tabla 2
Porcentaje de reactivos por operación cognitiva y tipo de conocimiento a evaluar

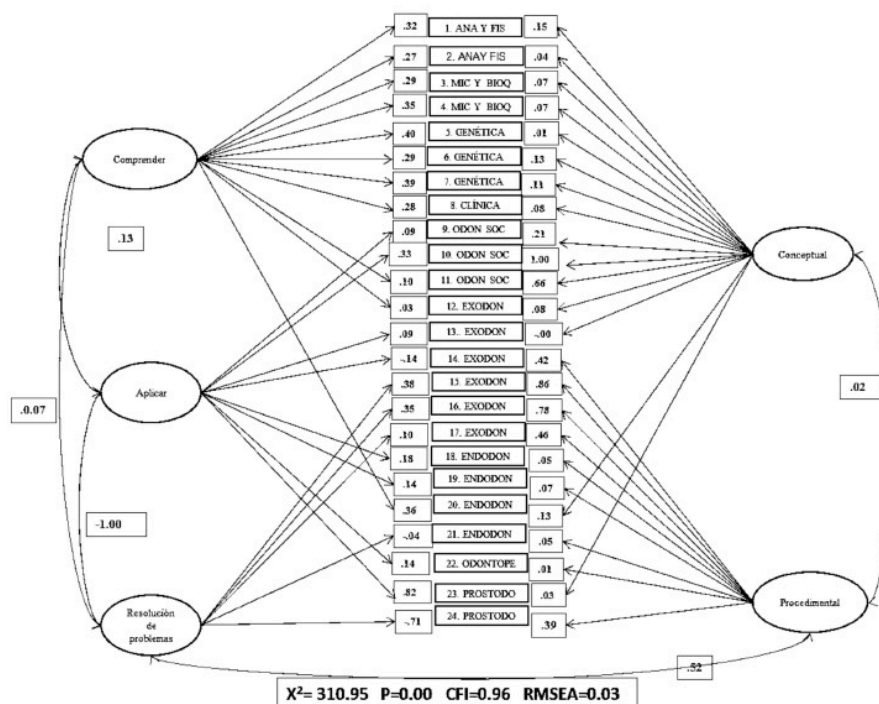


Figura 2
Parámetros estimados

Figura 2
Parámetros estimados

Se encontró que existe validez de constructo convergente entre las demandas cognitivas de comprensión y aplicación, cuyo peso factorial fue de .26. La siguiente relación que mostró tener validez convergente fue entre los procesos cognitivos de aplicación y solución de problemas, con carga factorial de 1.00.

Se observa validez de constructo divergente entre la demanda cognitiva de comprensión y solución de problemas, cuyo peso factorial fue .13. La relación entre los tipos de conocimiento conceptual y procedimental también muestra validez divergente ya que su peso factorial fue .031.

Al analizar los coeficientes de correlación estandarizados se encontró que en el factor procedimental, de las nueve variables relacionadas con él, todas presentan una relación alta. De forma contraria en el factor conceptual, de las 15 variables observadas que lo componen, cinco tienen una relación alta y corresponden con las subáreas de genética, odontología social, endodoncia y prostodoncia.

Respecto a los resultados en el factor comprender de las 11 variables relacionadas con él, siete de ellas tienen una relación media y las cuatro restantes muestran relación baja, siendo una variable de genética la que tiene una relación más alta, mientras que una variable de odontología social tiene la correlación más baja. En el factor aplicar se encontró que de ocho variables sólo una tiene una relación alta y corresponde con la temática de prostodoncia. De forma similar el factor de solución de problemas tiene cinco variables, de las cuales una muestra relación alta y las cuatro restantes tienen una relación no significativa.

Tabla 3
Índices de ajuste

Índices de ajuste	
Bentler-Bonett. Índices de ajuste normalizados	0.79
Bentler-Bonett. Índices de ajuste no normalizados	0.85
Índices de ajuste comparativo (CFI)	0.88
Bollen (IFI). Índices de ajuste	0.88
Mcdonald (MFI). Índices de ajuste	0.78
Joreskog-Sorbom GFI. Índices de ajuste	0.93
Joreskog-Sorbom AGFI. Índices de ajuste	0.9
Root Residual de media cuadrática (RMR)	0.01
RMR estandarizado	0.06
Root error de aproximación de media cuadrática (RMSEA)	0.05
90% de intervalo de confianza RMSEA	(.041, .053)
Coeficientes de confiabilidad	
Alfa de Cronbach	0.7
Coeficiente de confiabilidad RHO	0.7

Tabla 3
Índices de ajuste

Se observa en los resultados que seis variables contienen la mayor proporción de varianza explicada en el modelo, dos corresponden con las temáticas de odontología social y su varianza explicada es de 1.00 y .435 respectivamente; están relacionadas con la demanda cognitiva de comprensión y el tipo de conocimiento comprender. De forma semejante dos variables de exodoncia muestran varianza explicada de .903 y .771, respectivamente, en el modelo se relacionan con la demanda cognitiva de solución de problemas y el tipo de conocimiento procedimental. Por último, dos variables que representan los contenidos temáticos de prostodoncia tienen relación con la demanda cognitiva de solución y tipo de conocimiento procedimental.

Al analizar los coeficientes estandarizados del modelo, se observa una relación alta entre la demanda cognitiva comprender y el tipo de conocimiento conceptual, respecto a la relación entre el proceso cognitivo y el tipo de conocimiento de los temas de anatomía y fisiología, microbiología y bioquímica; genética e histoembriología y patología y farmacología presentan efecto medio entre todas las variables, de acuerdo con los resultados de las variables de estos temas.

Las dimensiones hipotetizadas tipo de conocimiento conceptual y el proceso cognitivo comprender tienen una correlación alta y significativa, los sustentantes del examen no tienen problema al enfrentarse a reactivos que evalúan conceptos e implican una operación mental de comprensión; sin embargo, esto no sucede con la relación de las dimensiones conceptual y aplicar, de acuerdo con los datos resultantes, la correlación es baja cuando los sustentantes del examen se enfrentan a reactivos que demandan la operación cognitiva de aplicación combinada con el tipo de conocimiento conceptual.

De forma similar, al ser examinados los sustentantes con reactivos que evalúan el tipo de conocimiento procedimental y al mismo tiempo llevan a cabo el proceso cognitivo de aplicación, se observa que la relación entre ambas dimensiones es alta y significativa; sin embargo, cuando la relación establece que los sustentantes deben responder reactivos que impliquen el tipo de conocimiento procedimental y realizar una operación cognitiva que involucre la resolución de problemas la correlación es baja.

Al analizar los contenidos del examen y su relación con los tipos de conocimiento y procesos cognitivos, se encuentra que el modelo está representado en su mayoría por reactivos que evalúan contenidos de genética e histopatología del área de formación básica, así como clínica básica del área de formación instrumental.

Los temas de genética e histopatología, así como el resto de los contenidos temáticos del área básica impartidos mayoritariamente en los primeros semestres de la carrera, se consideran fundamentales para el inicio del aprendizaje de los principios teóricos, prácticos y metodológicos que permitirán al futuro profesionista consolidar las bases hacia un aprendizaje gradual.

Respecto al área de formación instrumental el objetivo de evaluación consiste en la ejecución de tareas, actividades y funciones que exclusivamente puede realizar un odontólogo en relación con la especificidad de la atención de la salud bucal. Se consideran, asimismo, los contenidos que proporcionan los elementos metodológicos para la realización de las funciones de diagnóstico, determinación de planes de tratamiento y operación de programas de atención, tanto individual como comunitaria.

La subárea de clínica básica tiene por objetivo evaluar contenidos que demandan del sustentante la integración de conocimientos y habilidades que permiten identificar las manifestaciones clínicas más comunes en condiciones de salud, de emergencia y de urgencias, que afectan los órganos dentarios y tejidos de soporte. Se aplican los principios de la prevención, reconstrucción y conservación del sistema estomatognático; además de la farmacodinamia de los agentes anestésicos.

Este preámbulo se incluyó para facilitar la comprensión del modelo y sus efectos, ya que como se observa en la Figura 3.2 hipotéticamente se estableció que los tipos de conocimientos conceptuales tienen relación directa con la demanda cognitiva de comprensión y reactivos que evalúan temáticas de anatomía y fisiología, microbiología y bioquímica, genética e histoembriología y patología.

Los coeficientes estandarizados en el modelo demuestran que, en la dimensión mencionada de los nueve reactivos que la conforman, dos de ellos no presentan valores de correlación significativa, estos reactivos pertenecen a las subáreas anatomía y fisiología y genética e histoembriología, en resumen se observa que la hipótesis en la que se establece la relación entre reactivos conceptuales y procesos cognitivos de comprensión tienen una relación directa y significativa con el área que evalúa conocimientos básicos de la disciplina.

Contrario a los resultados de la relación entre el tipo de conocimiento conceptual, la demanda cognitiva de aplicación y reactivos del área de formación instrumental, donde la correlación es baja. Sin embargo, la misma demanda cognitiva de aplicación demuestra una correlación significativa con los reactivos de clínica básica, se comprueba la hipótesis que plantea la relación entre la aplicación de conocimientos para responder reactivos de solución de problemas.

En la hipótesis que establece la relación entre el tipo de conocimiento procedimental, el nivel cognitivo de aplicación y reactivos de la subárea de formación clínica demuestra una correlación significativa, lo anterior permite inferir que los sustentantes son capaces de aplicar técnicas y procedimientos, tales como establecer diagnósticos o elegir el tratamiento adecuado planteado en un caso clínico.

Por último, respecto al uso de técnicas y procedimientos combinado con el proceso cognitivo de solución de problemas no hay correlación. Pero si hay una relación directa y significativa con la subárea de clínica básica y la demanda cognitiva. Es evidente que dicha subárea permite evaluar a los sustentantes con diferentes niveles de complejidad y diferencia su pericia y experiencia de manera gradual.

DISCUSIÓN

Como ya se afirmó, el análisis del modelo revela que existe validez de constructo convergente entre el proceso cognitivo comprender y el tipo de conocimiento conceptual, estos resultados explican que existe una relación directa entre ambos procesos y su varianza explicada es de 7.03%.

En lo que respecta al proceso cognitivo de resolución de problemas y su relación con una variable que evalúa la aplicación de un tratamiento clínico, la varianza explicada en esta relación es de 22.53%.

De esta manera la evidencia recabada indica que los sustentantes responden de forma más homogénea cuando el tipo de conocimiento demandado para resolver los reactivos se relaciona con la operación cognitiva comprender.

Cabe aclarar que, aunque en el análisis factorial exploratorio se redujo el número de variables observadas y se seleccionaron aquellas que mostraban cargas factoriales significativas, los resultados del modelo estructural final demostraron una relación muy débil entre la operación cognitiva aplicar y el tipo de conocimiento conceptual, así como lo establecido en la relación del tipo de conocimiento procedimental y el proceso cognitivo de solución de problemas.

También es preciso mencionar que la selección de los reactivos finales que conformaron el modelo, fue resultado de varios análisis precedentes, tales como análisis factoriales exploratorios, por áreas de conocimiento del examen, si bien en modelo final contenía sólo una muestra muy pequeña de reactivos, se revisaron y analizaron los valores de más de 177.

A manera de conclusión se resalta la pertinencia de generar más evidencia sobre los constructos hipotetizados en los exámenes a gran escala, cuidando además de la pertinencia de los contenidos evaluados, los cuales se entienden como las áreas de conocimiento, los procesos cognitivos y tipos de conocimiento que son ejecutados por los futuros profesionistas.

Por último, este trabajo también deja abierta la pregunta sobre la adecuada aplicación y el rigor con el que se manejan los estándares de elaboración y diseño de los reactivos que conforman un instrumento de evaluación, así como el compromiso del personal docente que es sin duda órgano medular del diseño de exámenes y quien decide qué preguntas construir y con qué nivel de complejidad, para garantizar que éstos y no otros son los contenidos más adecuados y pertinentes para los objetivos planteados al decidir que se generará un nuevo instrumento de evaluación, cuyo propósito debe ser contribuir a la construcción del conocimiento acerca de los procesos de enseñanza y aprendizaje de un subsistema educativo.

Agradecimientos

Agradecemos a la Dirección de Ciencias de la Vida y la Conducta del Centro Nacional de Evaluación para la Educación Superior las facilidades otorgadas en la realización de este estudio.

REFERENCIAS

- Anastasi, A., & Urbina, S. (1998). *Test psicológicos*. México: Pretince Hall.
- Bentler, P. M. (2006). *EQS 6 structural equations program manual*. Encino, CA, US: Multivariate Software, Inc.
- Castañeda, S. (2004). *Educación, aprendizaje y cognición. Teoría en la práctica*. México: Manual Moderno.
- Cronbach, L. J., & Snow, R. E. (1969). *Individual differences in learning ability as a function of instructional variables: Final report*. Stanford, CA, US: US Department of Health, Education & Welfare.
- Delis, D. C., Jacobson, M., Bondi, M. W., Hamilton, J. M., & Salmon, D. P. (2003). The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: Lessons from memory assessment. *Journal of the International Neuropsychological Society*, 9(6), 936-946. <https://doi.org/10.1017/S1355617703960139>
- Embretson, S. (1985). *Test design: Developments in psychology and psychometrics*. Orlando, FL, US: Academic Press.

- Embretson, S. E. (1999). Generating items during testing: Psychometric issues and models. *Psychometrika*, 64(4), 407-433. <https://doi.org/10.1007/BF02294564>
- Kline, R. B. (1998). Software review: Software programs for structural equation modeling: AMOS, EQS, and LISREL. *Journal of Psychoeducational Assessment*, 16(4), 343-364. <https://doi.org/10.1177/073428299801600407>
- Long, J. S. (1983). *Confirmatory factor analysis: A preface to LISREL* (Vol. 33). London, UK: Sage Publications.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Tyler, D. D. (1969). Evidence of a phosphate-transporter system in the inner membrane of isolated mitochondria. *Biochemical Journal*, 111(5), pp. 665-678. <https://doi.org/10.1042/bj1110665>
- Whitely, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, 93(1), 179. <http://dx.doi.org/10.1037/0033-2909.93.1.179>
- Zimowski, M., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). *BILOG-MG 3: Item analysis and test scoring with binary logistic models*. Chicago, IL, US: Scientific Software.