



# Revista Electrónica de Psicología Iztacala



Universidad Nacional Autónoma de México

Vol. 25 No. 2

Junio de 2022

## CALIBRACIÓN DE REACTIVOS A TRAVÉS DEL MODELO RASCH EN LA EVALUACIÓN DE MÉDICOS INTERNOS DE PREGRADO

María Dolores Martínez Calvo<sup>1</sup>, Gerardo Muñoz Cortés<sup>2</sup>, María Elena Estrada Andrade<sup>3</sup>, Yurixhi Cancino Estrada<sup>4</sup>, Carlos Gómez Alonso<sup>5</sup> y Perla Patricia Abarca de Hoyos<sup>6</sup>

Instituto Mexicano del Seguro Social  
Michoacán, México

### RESUMEN

**Introducción:** El Médico Interno es pilar del sistema de salud. La evaluación de un alumno debe ser objetiva. El método Rasch relaciona la probabilidad de responder ítems correctamente, la diferencia entre el nivel en el constructo del individuo y la dificultad del ítem. **Objetivo:** Calibrar reactivos a través del Modelo Rasch en la evaluación de médicos internos de la Unidad de Medicina Familiar No. 80. **Material y métodos:** Estudio observacional, descriptivo. Realizado en Médicos Internos del Instituto Mexicano del Seguro Social en Morelia rotando durante los bimestres: enero-febrero, marzo-abril y julio-agosto del 2020. Muestreo por conveniencia, total 23 médicos internos. Se diseñó una primera evaluación, posteriormente se analizó, utilizando el modelo Rasch, con Ministep, después se desarrolló una evaluación final de reactivos calibrados, con estadígrafos de ajuste adecuados. Aprobado con el Número de Registro Institucional R-2019-1602-024. **Resultados:** Reactivos de los exámenes de conocimientos, inicial vs final: Infit 0.99 vs 1, Correlación de Pearson 1 vs 1 y Alfa de Cronbach 0.14 vs 0.61. **Conclusión:** se generaron exámenes calibrados a las habilidades de los médicos internos

**Palabras clave:** Evaluación objetiva, modelo Rasch, calibración de reactivos, Médico Interno.

<sup>1</sup> Instituto Mexicano del Seguro Social. Correo Electrónico: [marz\\_marcal@gmail.com](mailto:marz_marcal@gmail.com)

<sup>2</sup> Instituto Mexicano del Seguro Social. Correo Electrónico: [gerardomunozcortes@gmail.com](mailto:gerardomunozcortes@gmail.com)

<sup>3</sup> Universidad Michoacana de San Nicolás de Hidalgo. Correo Electrónico: [elencorreo@yahoo.com.mx](mailto:elencorreo@yahoo.com.mx)

<sup>4</sup> Instituto Mexicano del Seguro Social. Correo Electrónico: [yurixhi2@hotmail.com](mailto:yurixhi2@hotmail.com)

<sup>5</sup> Instituto Mexicano del Seguro Social. Correo Electrónico: [cga\\_0123@hotmail.com](mailto:cga_0123@hotmail.com)

<sup>6</sup> Instituto Mexicano del Seguro Social. Correo Electrónico: [perlabarca@gmail.com](mailto:perlabarca@gmail.com)

# CALIBRATION OF REAGENTS USING THE RASCH MODEL IN THE EVALUATION OF MEDICAL INTERNS

## ABSTRACT

**Introduction:** The medical intern is a pillar of the health system. The evaluation of a student must be objective. The Rasch method relates the probability of answering items correctly, the difference between the level in the individual's construct and the difficulty of the item. **Objective:** Calibrate reagents through the Rasch Model in the evaluation of medical interns of the Family Medicine Unit No. 80. **Material and methods:** Observational, descriptive study. Performed in medical interns of the Mexican Institute of Social Security in Morelia rotating during the two-month periods: January-February, March-April and July-August 2020. Convenience sampling, total 23 medical interns. A first evaluation was designed, later it was analyzed, using the Rasch model, with Ministep, then a final evaluation of calibrated reagents was developed, with adequate adjustment statistics. Approved with the Institutional Registration Number R-2019-1602-024. **Results:** Reagents of the knowledge tests, initial vs final: Infit 0.99 vs 1, Pearson's Correlation 1 vs 1 and Chronbach's Alpha 0.14 vs 0.61. **Conclusion:** tests calibrated to the skills of the internal doctors were generated  
**Keywords:** Objective evaluation, Rasch model, reagent calibration, Medical Intern.

## Internado Médico de Pregrado

De acuerdo con la Norma Oficial Mexicana NOM-234-SSA1-2003, llamada Utilización de campos clínicos para ciclos clínicos e internado de pregrado, se define al Internado de pregrado como el ciclo académico teórico-práctico que se realiza como parte de los planes de estudio de licenciatura en medicina, como una etapa que debe cubrirse previamente al servicio social, al examen profesional y al título respectivo (NOM-234-SSA1-2003).

El internado de pregrado tiene como fin desarrollar en los médicos internos de pregrado (MIP) un sentido de solidaridad con la sociedad general; contribuir con la formación integral y capacitación profesional de los futuros médicos; aplicar los conocimientos teóricos, técnicos y humanísticos, adquiridos por el interno de pregrado en su formación académica, relacionados con el diagnóstico, prevención, tratamiento y rehabilitación de los pacientes en cuya atención participan; fortalecer la formación académica, desarrollar valores y fortalecer la inserción al mercado de trabajo; así como al desarrollo comunitario, ya que es una actividad comprometida con los problemas de salud de la sociedad

(Acuerdo por el que se expiden los lineamientos generales para internos de pregrado, 2007).

La programación de rotaciones, horarios y actividades complementarias que realizan los médicos internos de pregrado durante ese año son reguladas por el Programa Operativo del Internado de Pregrado de cada hospital, el cual se basa en los recursos que posee la sede. El proceso educativo se centra en el médico interno, buscando estimular el autoaprendizaje, haciéndolo participe en la búsqueda de soluciones para un problema de salud, integrándolo a las sesiones académicas que se realizan en el centro hospitalario (Tapia, 2007).

Martínez (2016), menciona que es tan definitiva esta etapa que marca la diferencia en su desempeño profesional, por lo que realmente existe un antes y un después del Internado Médico de Pregrado. Los Médicos Internos de Pregrado son la evolución natural del médico, no son estudiantes, pero tampoco médicos, aún no saben sus debilidades ni fortalezas, aún se conocerán y descubrirán profesionalmente durante este año.

En el Instituto Mexicano del Seguro Social (IMSS), la Dirección de Prestaciones Médicas, mediante el documento llamado Procedimiento para la planeación, desarrollo, control y evaluación de los procesos educativos de pregrado: ciclos clínicos, internado y servicio social de las carreras del área de la salud regula el año de Internado médico de pregrado de todos los alumnos adscritos al instituto, coordinando la planeación de acuerdo a la disponibilidad de recursos, tanto materiales como humanos, para poder desarrollar de manera adecuada el Internado.

En este documento se define al Interno como “Alumno de la carrera de Licenciatura en Medicina que aprobó todas las materias de los ciclos básicos y clínicos establecidos en el plan de estudios correspondiente a la institución educativa de procedencia, el cual es considerado por la misma como candidato a realizar el internado con duración de un año y aceptado por el Instituto de acuerdo con los criterios del presente procedimiento” (Dirección de prestaciones médicas, 2014).

#### *Plan educativo*

En el Hospital General Regional No. 1 (HGR No. 1) del IMSS delegación Michoacán, los Médicos Internos de Pregrado provienen de dos instituciones educativas, la Universidad

Michoacana de San Nicolás de Hidalgo y la Universidad Vasco de Quiroga. El plan educativo se basa en el Manual del Médico Interno de Pregrado 2014-2015 que proporciona la Universidad Michoacana de San Nicolás de Hidalgo y en la parte académica rige a todos los Médicos Internos adscritos al IMSS en Michoacán, y en este caso en particular, a los Médicos Internos de HGR No. 1. El Manual divide el Programa del Internado Médico de Pregrado en Rotaciones, de las cuales cada una cuenta con un subprograma de autoaprendizaje y otro de educación en servicio. En la Rotación de Medicina Familiar que es la que nos concierne, encontramos que tiene como base 7 ejes centrales en las áreas de aprendizaje:

1. Atención primaria para la salud: Definición, objetivos, estrategias, equipos de salud, metodologías.
2. Nivel de atención médica: características generales de los niveles de atención, sistema de referencia y contrarreferencia de pacientes, sistemas locales de salud.
3. Programa Nacional de Salud: programas prioritarios, subprogramas específicos, formas de instrumentación.
4. Comunidad: técnicas de estudio, diagnóstico de salud.
5. Educación para la salud: diseño de programas de educación para la salud, técnicas didácticas auxiliares.
6. La familia dentro de la estructura ecológica y social: características de un grupo, conceptos, clasificación, la familia, definición, fases, ciclo de vida y funciones, fundamentos de la atención continua en Medicina Familiar
7. Medicina integral de pacientes y su familia en el primer nivel de atención: definición, complementos básicos de la medicina integral, engranaje de sus acciones, metodología para llevar una buena atención médica integral.

En cuanto a los temas que debe manejar el Interno en su rotación por Medicina Familiar se encuentran los siguientes: enfermedades transmisibles sexualmente, enfermedades prevenibles por vacunación, enfermedades laborales, alcoholismo y drogadicción, enfermedades crónico-degenerativas (diabetes, hipertensión arterial y obesidad, entre

otras), parasitosis, cólera, rabia, estudio de familia: familiograma y disfunciones familiares.

Y dentro de las habilidades y destrezas a desarrollar, hay cuatro que generan interés de acuerdo con la investigación que se lleva a cabo:

\*Detección de factores de riesgo: prevención, diagnóstico, tratamiento y rehabilitación de las enfermedades más frecuentes de la comunidad

\* Manejo correcto de los productos biológicos del esquema básico para niños de 0-4 años, preescolares, escolares, mujeres en edad fértil, adultos mayores;

\* Realizar atención materno infantil: control del embarazo normal, crecimiento y desarrollo, detección de embarazo de alto riesgo;

\*Control del niño sano: vigilar crecimiento y desarrollo, esquema de inmunizaciones, alimentación infantil. (Facultad de Ciencias Médicas y Biológicas "Dr. Ignacio Chávez", 2007).

#### Proceso educativo en los estudiantes de Medicina

El proceso educativo en Medicina ha evolucionado a lo largo del tiempo, Lafuente et. al (2007) refieren que desde 1950 se logran identificar tres periodos:

1. Un primer periodo de planificación por objetivos, inspirado por los métodos de gestión que se han impuesto a partir de los años cincuenta.
2. Un segundo periodo a partir de los años setenta, caracterizado por la búsqueda de la pertinencia con el fin de conseguir un médico en concordancia con la situación sanitaria local, con la motivación para identificar las necesidades cualitativas y cuantitativas.
3. Un tercer periodo, a partir de los noventa, caracterizado por la búsqueda del impacto de la educación médica sobre la salud de la población, lo que implica un proceso tendente a transformar la Facultad de Medicina en un "partenaire" explícito del desarrollo sanitario.

La educación médica requiere centrarse en las particularidades inherentes a las distintas ramas de las ciencias médicas que implica la unidad de lo docente y lo científico, por

cuanto se estudia la ciencia en desarrollo a fin de alcanzar la apropiación de sus métodos, mediante el despegamiento didáctico, lo que permite la constitución de los referentes profesionales sobre bases sólidas y el desarrollo de capacidades de autopreparación permanente durante toda la vida profesional.

El proceso de enseñanza-aprendizaje en las disciplinas clínicas se caracteriza por la interrelación entre los objetivos, contenidos, métodos, medios y la evaluación, donde los objetos ocupan el papel rector de lo que se pretende lograr en el estudiante, proceso docente que se organiza y desarrolla a través de la integración docencia-servicio-investigación; pero es en el escenario real donde se alcanzan los objetivos de la formación, por ser el medio idóneo para establecer la relación teórico-práctica, expresión máxima para el aprendizaje y consolidación de conocimientos y habilidades, así como para desarrollar las capacidades para recoger datos, diagnosticar, tratar y tomar decisiones de acuerdo a las características de los problemas que se le presenten (UNAM, 2002).

#### Hipertensión arterial sistémica

De acuerdo con Tagle (2018) “la hipertensión arterial es uno de los principales factores de riesgo de las enfermedades cardiovasculares”. Es por ello por lo que se considera un tema prioritario de salud, que los estudiantes de Medicina y los médicos deben dominar. El diagnóstico de Hipertensión arterial sistémica (HAS) se realiza cuando se detectan cifras de presión arterial mayores o iguales a 140/90 mm Hg en varias determinaciones, además de tomar en cuenta que el diagnóstico se confirme, realizando todas las mediciones de la misma manera y con el mismo instrumento; se deben considerar las cifras limítrofes entre 135-139 mm Hg de la presión sistólica y entre 85-89 mm Hg de la presión diastólica, como un indicador para buscar intencionadamente la presencia de HAS (Rosas-Peralta et al., 2016).

La hipertensión arterial causa diversas complicaciones en diferentes órganos: corazón, riñón, ojo y vasos sanguíneos, entre otros. Por lo tanto, conocer las acciones que se deben realizar para detectar estas complicaciones es mandatorio para el médico, con el fin de que su tratamiento se instaure de manera oportuna.

El control del paciente hipertenso es la meta a la que se debe aspirar, el paciente debe mantener cifras de tensión arterial menores a 140/90 mm Hg, en pacientes jóvenes y de mediana edad deben ser inferiores a 130/85 mm Hg. En pacientes diabéticos menor a 130/85 mm Hg. En pacientes que presentan daño renal se debe mantener menor a 125/75 mm Hg (Rosas-Peralta et al., 2006).

### Evaluación

La evaluación del médico interno de pregrado se debe considerar de manera integral, tomar en cuenta qué evaluar, así como cuándo y cómo evaluar, de esta manera no sólo se tomarán en cuenta las pruebas aplicadas al alumno, sino también la dinámica didáctica de las clases en las que estuvo el evaluado.

Partiendo del hecho de que la evaluación debe ser integral, debemos basarnos en las tres variedades de evaluación: diagnóstica, formativa y sumativa.

#### *Evaluación diagnóstica.*

Orozco-Jutorán, (2006), refiere que la evaluación diagnóstica “se centra en el tipo y nivel de conocimientos que tienen los alumnos antes de iniciar un curso o asignatura”. En este sentido sugiere que se hagan dos pruebas de seguimiento, al inicio y al final del curso para comparar los conocimientos adquiridos por los estudiantes.

#### *Evaluación formativa.*

La Agencia de Calidad de la Educación [ACE] (2016) define la evaluación formativa como “un proceso en el cual profesores y estudiantes comparten metas de aprendizaje y evalúan constantemente sus avances en relación con estos objetivos”. El fin de la evaluación formativa es establecer la mejor manera de continuar el proceso didáctico de acuerdo con las necesidades del grupo. Se rige mediante tres principios: hacia dónde se quiere llegar, dónde se está y cómo hacer para alcanzar la meta (ACE, 2016).

#### Evaluación sumativa.

Este tipo de evaluación “Tiene por objetivo establecer balances fiables de los resultados obtenidos al final de un proceso de enseñanza-aprendizaje. Pone el acento en la recogida de información y en la elaboración de instrumentos que posibiliten medidas fiables de los conocimientos a evaluar” (Rosales, 2014 p. 4).

### Instrumentos de medición para la evaluación

De acuerdo con Sampieri (2014), en la investigación se dispone de varios instrumentos que miden las variables de interés y a veces se combinan varias técnicas de recolección de datos. Dentro de los instrumentos para medir las variables encontramos los cuestionarios y las escalas para medir las actitudes.

Los primeros son más utilizados en fenómenos sociales y consisten en un conjunto de preguntas para medir una o más variables en congruencia con el planteamiento del problema y la hipótesis; dependiendo el problema de investigación y la necesidad que se busque cubrir con este instrumento las preguntas pueden ser abiertas o cerradas. Los cuestionarios se pueden aplicar de manera autoadministrada o por medio de una entrevista.

Las escalas para medir la actitud se enfocan en medir la dirección e intensidad de las actitudes del evaluado, las cuales son un indicador de la conducta. Dentro de estas escalas las más conocidas son el Escalamiento de Likert, el Diferencial semántico y la Escala de Guttman

Existen otros métodos cuantitativos para la recolección de datos (Sampieri, 2014):

1. Análisis de contenido cuantitativo
2. Observación
3. Pruebas estandarizadas e inventarios
4. Datos secundarios (recolectados por otros investigadores)
5. Instrumentos mecánicos o electrónicos
6. Instrumentos y procedimientos específicos propios de cada disciplina

Según Ferreirós (2018), para la evaluación del conocimiento las herramientas a utilizar pueden ser exámenes o pruebas de aptitud. Y dentro de los exámenes podemos utilizar los siguientes:

- Preguntas de elección múltiple o exámenes tipo test: miden esencialmente conocimientos. Son muy eficientes, fiables, y son útiles para pruebas con elevado número de examinados



- Evaluación con preguntas abiertas: tiene un perfil similar a las preguntas de opción múltiple y presentan una mayor competencia para valorar la capacidad de toma de decisiones y resolución de problemas.
- Examen escrito u oral: valora la integración de conocimientos y la toma de decisiones. Es más fiable cuando es estructurado.

### Formato de ítems de opción múltiple

De acuerdo con Moreno (2004) los ítems de elección múltiple exigen a quien los elabora un mayor esfuerzo y pericia que aquellos más abiertos que sólo plantean la pregunta, aunque hoy por hoy tienen la ventaja de una evaluación más fiable.

Para que una pregunta de un examen sea aceptable debe cumplir dos criterios básicos: incluir contenido relevante y estar correctamente estructurados, esto se logra evitando defectos que logren beneficiar a candidatos que saben rendir exámenes y evitando dificultades irrelevantes. Lo cual permite que las preguntas del examen generen puntajes válidos. En este formato de ítems entran dos modalidades: aquellas que requieren que el alumno indique todas las respuestas que son correctas y aquellas que requieren que el alumno indique una sola respuesta (Case y Swanson, 2006).

### Teoría de Respuesta al Ítem

De acuerdo con Muñiz (2010), la Teoría clásica de los test (TCT) presenta dos limitaciones: la primera que las mediciones no resultan invariantes respecto al instrumento utilizado y la segunda es que las propiedades psicométricas importantes de los tests, tales como la dificultad de los ítems, o la fiabilidad del test, estaban en función del tipo de personas utilizadas para calcularlas, lo cual es inadmisibles desde el punto de vista de una medición rigurosa. Esas dificultades se tratan de resolver con la Teoría de Respuesta al Ítem (TRI).

Esta denominación agrupa líneas de investigación independientes iniciadas por Rasch y Birnbaum con un factor en común que establece una relación entre el comportamiento de un sujeto frente a un ítem y el rasgo responsable de esta conducta (Attorresi, Lozzia, Abal, Galibert, y Aguerri, 2009).

El objetivo sustancial de esta teoría es la construcción de instrumentos de medición con propiedades invariantes entre poblaciones; aquí se toma al ítem como unidad de análisis y se modeliza directamente sobre el rasgo latente; esto significa que la TRI permite mediciones que no varíen más allá de los ítems que componen el instrumento (Attorresi, 2009).

### Modelo Rasch

En el ámbito educativo, se han utilizado pruebas estandarizadas para medir constructos complejos, las cuales se basan en la Teoría Clásica de los Test, ya que los puntajes que se obtienen mediante esas pruebas no proveen información acerca de qué puede o no hacer el estudiante. El significado que se otorga a las puntuaciones de una prueba bajo el modelo TCT se define en torno a la posición relativa que ocupa un examinado particular en relación con el grupo total de examinados que tomó el test; en un contexto educativo es además necesario establecer la posición relativa del examinado, ser capaz de utilizar las puntuaciones de las pruebas para determinar a nivel absoluto las fortalezas y debilidades del individuo en torno al constructo que interesa medir, lo que se conoce como interpretaciones referidas a criterios (Rojas y Montero, 2012).

El modelo Rasch es un recurso que analiza la calidad técnica de ítems que forman parte de una prueba, su fórmula matemática relaciona la probabilidad de responder el ítem correctamente, la diferencia entre el nivel en el constructo o atributo del individuo y la dificultad del ítem (Rasch, 1960).

El proceso lógico que rige al modelo Rasch es sencillo y parsimonioso y puede determinarse de la siguiente manera:

El modelo postula que un reactivo eficiente sólo debe ser contestado correctamente por aquellos sustentantes que posean la habilidad requerida por las demandas cognitivas del reactivo. A la inversa, un reactivo eficiente no debe ser contestado correctamente por aquellos sustentantes que no posean la habilidad requerida para responderlo (Montesinos, 2008).

De acuerdo con López (1998), el análisis de Rasch estudia qué tanto se acercan los datos medidos al modelo. Haciendo una recapitulación, se puede conjeturar lo siguiente para el análisis de Rasch:

- La escala que se maneja debe ser lineal
- La escala debe permitir realizar medidas de reactivos y personas referidas a un origen común
- La variable que se mide representa la diferencia entre la habilidad de la persona y la dificultad del instrumento
- La variable es unidimensional, ordenada e inclusiva, de acuerdo con el modelo del Escalograma de Guttman
- Las medidas son estocásticas, con una probabilidad de ocurrencia tanto en aciertos como en errores.

Por lo tanto, este estudio buscó aplicar la calibración de reactivos a través del Modelo Rasch en la evaluación de los Médicos Internos de Pregrado de la Unidad de Medicina Familiar Número 80.

### Material y Métodos

Se realizó un estudio observacional, transversal, descriptivo.

La población de estudio fueron los Médicos Internos de Pregrado (MIP) que se encontraban realizando su Internado rotatorio de pregrado en el Instituto Mexicano del Seguro Social (IMSS), adscritos al Hospital General Regional 1 (HGR1), de Morelia Michoacán, durante los bimestres enero-febrero, marzo-abril y julio-agosto del 2020.

Para el tamaño de la muestra se utilizó la fórmula de población finita:

$$n = N / (1 + (N * e^{**2}))$$

En donde N = 74 (total de MIP); e = 0.12 (error a tolerar). Al sustituir queda una n = 35.

El muestreo se realizó de manera no probabilística, por conveniencia. Se incluyeron a los MIP que se encontraban rotando en el servicio de Medicina Familiar, durante los bimestres enero-febrero, marzo-abril y julio-agosto del 2020.

Como criterios de selección se incluyeron a los MIP de sexo masculino o femenino, de cualquier edad, que aceptaron participar y que no tuvieron guardia el día previo a la aplicación de los exámenes. Se descartaron a los MIP que declinaron participar, a los

que terminaron su rotación por el servicio de Medicina Familiar antes de terminar el estudio, a los que cursaron con una incapacidad de más del 20% de asistencia al servicio. Para la calibración de reactivos se utilizó el Modelo Rasch, tomando en cuenta los siguientes estadígrafos de ajuste: Infit MNSQ entre 0.8 y 1.3; Alfa de Cronbach  $\geq$  6.0; Correlación de Pearson entre 0.9 y 1.10 para las personas y entre -0.9 y -1.10 para reactivos, en el caso de que los reactivos cumplieran con los tres estadígrafos se consideraron calibrados.

Previa autorización del comité de ética en investigación el protocolo fue aprobado por el Comité Local de Investigación en Salud 1602 con el Número de Registro Institucional R-2019-1602-024. Se hizo la invitación a los Médicos Internos de Pregrado, para que participaran en el estudio. A aquellos alumnos que decidieron participar y firmaron el consentimiento informado, se les aplicó un examen de conocimientos con respuesta de opción múltiple, acerca del tema de Hipertensión Arterial Sistémica.

Se realizó un examen diagnóstico al inicio del bimestre con una evaluación por tema prioritario que constó de 15 reactivos, que no estaban calibrados. Posteriormente se calificaron las evaluaciones y se introdujeron los datos a la aplicación Winstep para analizar los estadígrafos de ajuste de cada reactivo y se determinó qué reactivos se encontraban calibrados y cuáles no.

Se formuló un nuevo examen del tema, corrigiendo los reactivos no calibrados y se aplicó a los Médicos Internos de pregrado al final del bimestre en el que se encontraban rotando por la UMF 80. Después se calificó la evaluación y se analizaron los reactivos en la aplicación Winstep para determinar si los reactivos que se corrigieron ya se encontraban calibrados.

Tanto la evaluación diagnóstica como la evaluación final se realizaron en las aulas de la UMF 80 en la ciudad de Morelia, Michoacán. Se les dio de forma impresa y contaron con una hora para resolverla.

Todos los cálculos se realizaron con el paquete estadístico SPSS versión 22.0 para Windows. Los datos se presentan en frecuencias con sus respectivos porcentajes, se usó media (M) con su desviación estándar (DE); se midió la fiabilidad de los instrumentos de evaluación mediante alfa de Cronbach. El instrumento de evaluación objetiva se calibró mediante el modelo Rasch, a través del programa Winstep.

### Resultados

Se elaboró un examen sobre el tema de Hipertensión Arterial Sistémica, que constó de 20 reactivos, el cual fue aplicado a los MIP que se encontraban rotando en el servicio de Medicina familiar en la UMF 80. Se incluyeron 13 médicos internos de pregrado que cumplieron con los criterios de selección, 69.24% mujeres y 30.76% hombres, con edad de  $23 \pm 1$  año.

TABLE 3.1 ANALISIS DE EXAMEN DE CONOCIMIENTOS ZOU690WS.TXT Apr 28 2021 19:38  
 INPUT: 13 Alumno 25 Reactivo REPORTED: 13 Alumno 20 Reactivo 2 CATS MINISTEP 3.90.3

SUMMARY OF 13 MEASURED Alumno

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	8.2	20.0	-.22	.57	1.00	.0	.97	.0
P.SD	2.9	.0	.89	.02	.23	.9	.31	.7
S.SD	3.0	.0	.93	.02	.24	.9	.33	.7
MAX.	13.0	20.0	1.25	.60	1.38	1.4	1.72	1.4
MIN.	5.0	20.0	-1.23	.54	.66	-1.6	.55	-1.2
REAL RMSE	.60	TRUE SD	.66	SEPARATION	1.10	Alumno RELIABILITY	.55	
MODEL RMSE	.57	TRUE SD	.68	SEPARATION	1.20	Alumno RELIABILITY	.59	
S.E. OF Alumno MEAN = .26								

Alumno RAW SCORE-TO-MEASURE CORRELATION = 1.00  
 CRONBACH ALPHA (KR-20) Alumno RAW SCORE "TEST" RELIABILITY = .61 SEM = 1.82

Figura 1. Análisis estadístico por Winstep de los alumnos.

El programa Winstep arrojó los resultados estadísticos de los 13 alumnos a los que se les aplicó el examen, como se observa en la figura 1, se obtuvo una media de 8.2 respuestas correctas de los 20 reactivos, además que el promedio en el Infit MNSQ se encuentra en 1 lo cual es un estadígrafo de calibración adecuado. El alfa de Cronbach se encuentra en 0.61, cuando de manera inicial, previo a la calibración, se encontraba en 0.14. En este caso la correlación de Pearson se mantuvo en 1, siendo adecuada a la calibración.



*Evaluación de Hipertensión Arterial Sistémica*

<i>Estadísticos</i>	Iniciales	Finales
<i>Infit MNSQ</i>	0.99	1
<i>Correlación de Pearson</i>	1	1
<i>Alfa de Cronbach</i>	0.14	0.61

Tabla I. Comparativa entre los estadísticos iniciales y finales de la evaluación sobre Hipertensión Arterial. Reactivos calibrados con valores de Infit MNSQ entre 0-8 y 1.3, correlación de Pearson entre 0.9 y 1.10, Alfa de Cronbach  $\geq 6$ .

El examen inicial se analizó utilizando el programa Ministep, que se basa en el Modelo Rasch para determinar las características de cada reactivo, posteriormente se ajustaron las preguntas y se aplicó el nuevo examen a los mismos MIP, una vez que se obtuvieron los resultados se volvió a analizar el examen con el programa Ministep, el cual nos arrojó los nuevos resultados de los estadísticos. La comparativa de valores se muestra en la tabla I.

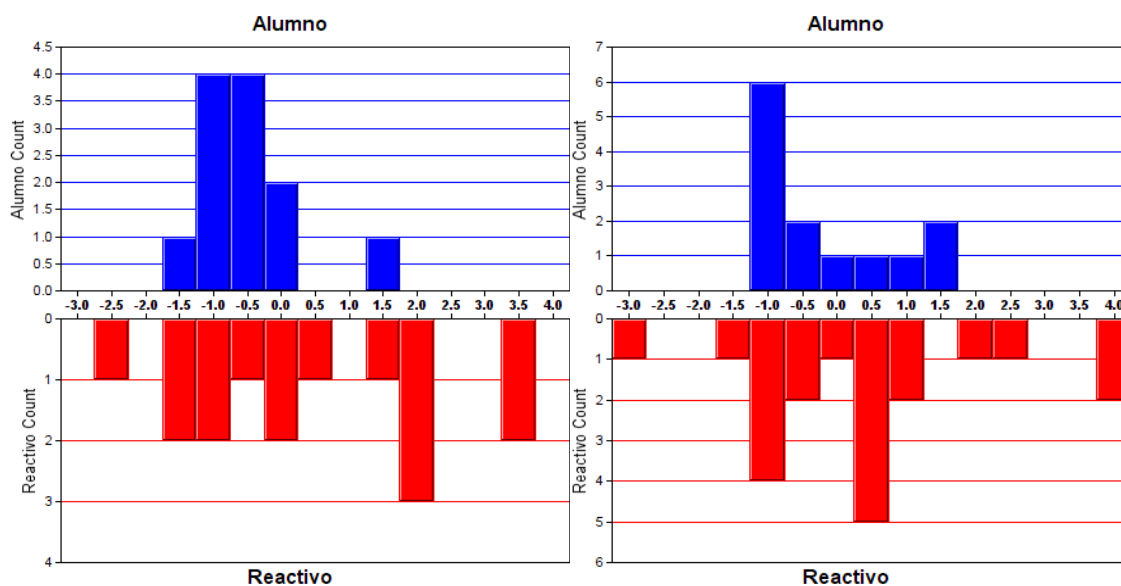


Figura 3. Análisis inicial y final del examen de Hipertensión Arterial Sistémica.

En la figura 3 se puede observar de manera más puntual como se calibraron los reactivos al conocimiento de los alumnos, esto se refleja en las gráficas en el apartado de reactivos, con los reactivos que van desde -3 a 4, y el grueso de los reactivos se

concentran en el rango -1.5 a 1; de igual manera los alumnos entran en concordancia con los reactivos al pasar de estar entre -1.5 a 1.5 a -1 a 1.5.

## Discusión

Actualmente la mayor parte de las evaluaciones se basan en la Teoría Clásica de los Test, las cuales son deficientes a la hora de hacer una evaluación imparcial y que refleje el grado de aprovechamiento de cada alumno. Esta investigación buscó realizar la calibración de reactivos mediante el modelo Rasch, para llevar a cabo una evaluación objetiva.

El primer objetivo específico que se tenía en mente era desarrollar instrumentos de evaluación calibrados que midan el tema de Hipertensión Arterial Sistémica. Se llevó a cabo la calibración de reactivos en el examen sobre el tema mencionado, en dichos reactivos se incrementó el valor del Alfa de Cronbach de 0.14 a 0.61, mientras que el valor del Infit MNSQ fue de 1; como menciona Pérez-López (2020) en su trabajo el cuestionario PIPK mostró adecuadas características psicométricas (alfa = 0.89) y buen ajuste del modelo de Rasch. Por otro lado, Azpilicueta (2019) encontró los siguientes resultados en su investigación al llevar a cabo el análisis de la CDS, el ajuste de los ítems varió entre  $-0.83 \leq \delta_i \leq 1.12$  ( $M = 0.00$ ;  $DE = 0.49$ ), los valores Infit variaron entre 0.82 y 1.41 ( $M = 1.02$ ;  $DE = 0.18$ ) y los Outfit entre 0.66 a 1.58 ( $M = 1.04$ ;  $DE = 0.28$ ), esto difiere en nuestra investigación, al notar que incluyeron también los valores de Outfit como criterio de calibración. De igual manera Vargas-Porras (2018) refiere en sus resultados que los 6 indicadores del resultado Nursing Outcomes Classification (2211) presentaron un alfa de Cronbach (0.90), valores Outfit entre 0.74 a 1.60 en este caso los valores, en este caso no se incluyó el valor de Infit como criterio, el cual sí fue tomado en cuenta en nuestros criterios.

El segundo objetivo que se pretendió alcanzar con esta investigación era realizar una evaluación diagnóstica y sumativa a los Médicos Internos de Pregrado acerca de los temas de Medicina Familiar, el cual se logró al llevar a cabo la evaluación diagnóstica inicial con un instrumento no calibrado, posterior a la calibración se realizó una segunda evaluación sumativa, la cual permitió establecer un plan didáctico para cubrir los temas prioritarios de salud en Medicina Familiar, con los valores de alfa de Cronbach



previamente mencionados, lo cual es semejante a los resultados obtenidos por Martínez González (2017), quien menciona que un 7.2% de los reactivos fueron eliminados en el análisis posterior, lo cual es apropiado para un examen de altas consecuencias y mejor si es un examen diagnóstico y formativo, como en nuestro caso, y el valor del alfa de Cronbach es 0.89, que corresponde a una alta consistencia interna. En concordancia con nuestros resultados Tinoco Rivera (2016) refiere en su investigación, que la evaluación diagnóstica, formativa y sumativa permite hoy en día no sólo sustentar un proceso de enseñanza-aprendizaje integral; sino evidenciar los resultados y logros alcanzados por los estudiantes y por las instituciones educativas, lo cual encontramos nosotros de igual manera ya que esa evaluación nos indica el nivel de formación universitaria de los Médicos Internos de Pregrado. Martínez, concluye en su trabajo que, los instrumentos de evaluación diagnóstica y sumativa cumplieron plenamente con su función ya que permitieron contrastar los ideales que tenían los alumnos antes y después de las dos sesiones de aprendizaje recibidas, mostrando en dos de tres casos un avance considerable; esto concuerda con nuestros resultados al notar una mejora en las calificaciones finales, después de llevar a cabo la calibración del instrumento y de que los Médicos Internos rotaran por el servicio de Medicina Familiar.

Durante la realización de la presente investigación se encontraron ciertas limitantes a tomar en consideración, la primera el hecho de que iniciamos un periodo de pandemia, causada por la aparición del COVID-19 lo cual modificó los tiempos de rotación de los Médicos Internos de pregrado en el servicio, así como la disponibilidad de los mismo para llevar a cabo las evaluaciones; otra de las limitantes es que la calibración de reactivos se llevó a cabo en 10 instrumentos y no uno solo, lo cual mostró diferentes resultados para cada uno. Otra limitante fue que el tiempo para resolver los exámenes varió entre las evaluaciones diagnósticas iniciales y las evaluaciones con instrumentos calibrados, debido a los cambios de horarios que hubo en la Unidad de Medicina Familiar. A pesar de las limitantes se tuvieron resultados satisfactorios.

## Conclusiones

La calibración de reactivos a través del modelo Rasch se llevó a cabo, lo cual nos permitió obtener resultados confiables que nos indican que la evaluación de los médicos Internos

se dio de manera objetiva. El alfa de Cronbach aumentó en promedio de la evaluación diagnóstica con instrumentos no calibrados a la evaluación sumativa que se realizó con instrumentos calibrados y sirvió para mejorar la calidad de la educación que se imparte al rotar por el servicio de Medicina Familiar.

## Referencias Bibliográficas

- Acuerdo por el que se expiden los lineamientos generales para internos de pregrado, Gaceta Oficial del Distrito Federal No. 238, Distrito Federal, México, 21 de diciembre 2007.
- Agencia de Calidad de la Educación (2016). *Guía de evaluación formativa*. Santiago de Chile, Chile.
- Attorresi, H. F., Lozzia, G. S., Abal, F. J., Galibert, M. S., y Aguerri, M. E. (2009). Teoría de Respuesta al Ítem. Conceptos básicos y aplicaciones para la medición de constructos psicológicos. *Revista Argentina de Clínica Psicológica*, XVIII (2), 179-188.
- Azpilicueta, A. E., Cupani, M., Ghío, B., Morán, V. E., y Garrido, S. J. (2019). Adaptación mediante el modelo de Rasch de tres medidas para estimar la decisión e indecisión de carrera y la ansiedad decisional.
- Case, S. M., y Swanson, D. B. (2006). *Cómo elaborar preguntas para evaluaciones escritas en el área de ciencias básicas y clínicas*. Filadelfia, Pensilvania, Estados Unidos: National Board of Medical Examiners.
- Dirección de prestaciones médicas. (2014). Procedimiento para la planeación, desarrollo, control y evaluación de los procesos educativos de pregrado: ciclos clínicos, internado y servicio social de las carreras del área de la salud. México, D. F.: IMSS.
- Facultad de Ciencias Médicas y Biológicas "Dr. Ignacio Chávez" (2007). *Manual de Médico Interno de Pregrado 2014-2015*. Morelia, Michoacán, México. Departamento de Educación Médica de la UMSNH.
- Ferreirós-Martínez, R., Hidalgo, B., Lorenzo, M., Carbonell, R., Alonso-Cerezo, M. (2018). Herramientas para la evaluación de las competencias profesionales. *Rev Lab Clin*, 11(2), 93-96. Madrid, España. Elsevier España.
- Lafuente, J., Escanero, J., Manso, J., Mora, S., Miranda, T., Castillo, M. (2007). Curricular design by competences in medical education: impact on the professional training. *Educ. méd.* 10(2), 86-92. Recuperado de

[http://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S1575-18132007000300004&lng=es](http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1575-18132007000300004&lng=es).

- López, A. T. (1998). *Análisis de Rasch para todos*. San Luis Potosí, México: Instituto de Evaluación e Ingeniería avanzada.
- Martínez, A. C. G., Vilchis, P. T. R., y Sánchez, A. E. R. La evaluación diagnóstica y sumativa para la mejora del aprendizaje de la historia.
- Martínez González, A., Lifshitz Guinzberg, A., Trejo Mejía, J. A., Torruco García, U., Fortoul van der Goes, T. I., Flores-Hernández, F., ... y Sánchez-Mendiola, M. (2017). Evaluación diagnóstica y formativa de competencias en estudiantes de medicina a su ingreso al internado médico de pregrado. *Gac Med Mex*, 153(1), 6-15.
- Martínez, M. (2016). Prólogo a la primera edición. En B. Valente (2ª Ed.), *Manual del Médico Interno de Pregrado* (pp. xv). Ciudad de México, México. Editorial Intersistemas.
- Montesinos, M. J. (2008). *El Análisis de Reactivos con el Modelo Rasch. Manual Técnico A*. México, D. F.: Instituto Nacional para la Evaluación de la Educación.
- Moreno, R., Martínez, R. J., Muñiz, J. (2004). Directrices para la construcción de ítems de elección múltiple. *Psicothema*, 16(3), 490-497. Oviedo, España. Universidad de Oviedo.
- Muñiz, J. (2010). Las teorías de los test: teoría clásica y teoría de respuesta a los ítems. *Papeles del psicólogo*, 31(1), 57-66. Madrid, España. Consejo General de Colegios Oficiales de Psicólogos.
- Norma Oficial Mexicana NOM-234-SSA1-2003, Diario Oficial de la Federación, Distrito Federal, México, 20 enero de 2003.
- Orozco-Jutarán, M., (2006). La evaluación diagnóstica, formativa y sumativa en la enseñanza de la traducción, en: Varela, M. J. (ed.) *La evaluación en los estudios de traducción e interpretación*. Sevilla: Bienza. p. 47-68 ISBN: 978-84-933962-8-2.
- Pérez-López, C., López-Franco, M. D., Comino-Sanz, I. M., y Pancorbo-Hidalgo, P. L. (2020). Validación del cuestionario de Conocimientos sobre prevención de lesiones por presión en estudiantes de enfermería: análisis Rasch. *Enfermería Clínica*.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research

- Rojas, S., Montero, E. (2012). El modelo Rasch: una herramienta esencial para la evaluación educativa de gran escala. Cuarto Informe del Estado de la educación
- Rosales, M. (noviembre, 2014). Proceso evaluativo: evaluación sumativa, evaluación formativa y Assessment su impacto en la educación actual. Trabajo presentado en el *Congreso Iberoamericano de ciencia, tecnología, innovación y educación*, Buenos Aires, Argentina.
- Rosas-Peralta, Martín, y Palomo-Piñón, Silvia, y Borrayo-Sánchez, Gabriela, y Madrid-Miller, Alejandra, y Almeida-Gutiérrez, Eduardo, y Galván-Oseguera, Héctor, y Magaña-Serrano, José Antonio, y Saturno-Chiu, Guillermo, y Ramírez-Arias, Erick, y Santos-Martínez, Efrén, y Díaz-Díaz, Enrique, y Salgado-Pastor, Selene Janette, y Morales-Mora, Gerardo, y Medina-Concebida, Luz Elena, y Mejía Rodríguez, Oliva, y Pérez-Ruiz, Claudia Elsa (2016). Consenso de Hipertensión Arterial Sistémica en México. *Revista Médica del Instituto Mexicano del Seguro Social*, 54(1),6-51. [fecha de Consulta 28 de Abril de 2021]. ISSN: 0443-5117. Disponible en:  
<https://www.redalyc.org/articulo.oa?id=457746536002>
- Sampieri, R. H. (2014). *Metodología de la Investigación*. Ciudad de México: Mc Graw Hill
- Tagle, R. (2018). Diagnóstico de hipertensión arterial. *Revista Médica Clínica Las Condes*, 29(1), 12-20.
- Tapia, R., Núñez, R., Salas, R., y Rodríguez-Orozco, A. (2007) El internado médico de pregrado y las competencias clínicas. México en el contexto latinoamericano, *Educ Med Super* 21 (4), 1-10.
- Tinoco Rivera, N. A. (2016). Incidencia de la evaluación diagnóstica, formativa y sumativa en los resultados de aprendizajes en la asignatura de estudios sociales.
- Vargas-Porras, C., Villamizar-Carvajal, B., y Roa-Díaz, Z. M. (2018). Validación de un instrumento en una unidad neonatal mediante el análisis Rasch. *UstaSalud*, 17, 15-22.
- Universidad Nacional Autónoma de México (2002). Curso de actualización en estrategias de enseñanza y evaluación del área clínica. México, DF: UNAM