



CODEIC

COORDINACIÓN DE DESARROLLO
EDUCATIVO E INNOVACIÓN
CURRICULAR UNAM

El examen de ingreso a las Licenciaturas de la UNAM

Dr. Melchor Sánchez Mendiola

Coordinador

melchorsm@unam.mx

ACUERDO QUE REORGANIZA LAS FUNCIONES Y ESTRUCTURA DE LA SECRETARÍA GENERAL DE LA UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

TERCERO.- Para el cumplimiento adecuado de sus funciones, el apoyo a los cuerpos colegiados y la coordinación de las dependencias adscritas a ella, la Secretaría General contará con las siguientes dependencias:

- I. Coordinación de Estudios de Posgrado;
- II. Coordinación de Desarrollo Educativo e Innovación Curricular;
- III. Coordinación de Vinculación con el Consejo Universitario;
- IV. Dirección General de Administración Escolar;
- V. Dirección General de Asuntos del Personal Académico;
- VI. Dirección General de Bibliotecas, y
- VII. Dirección General de Incorporación y Revalidación de Estudios.

30 de noviembre de 2015

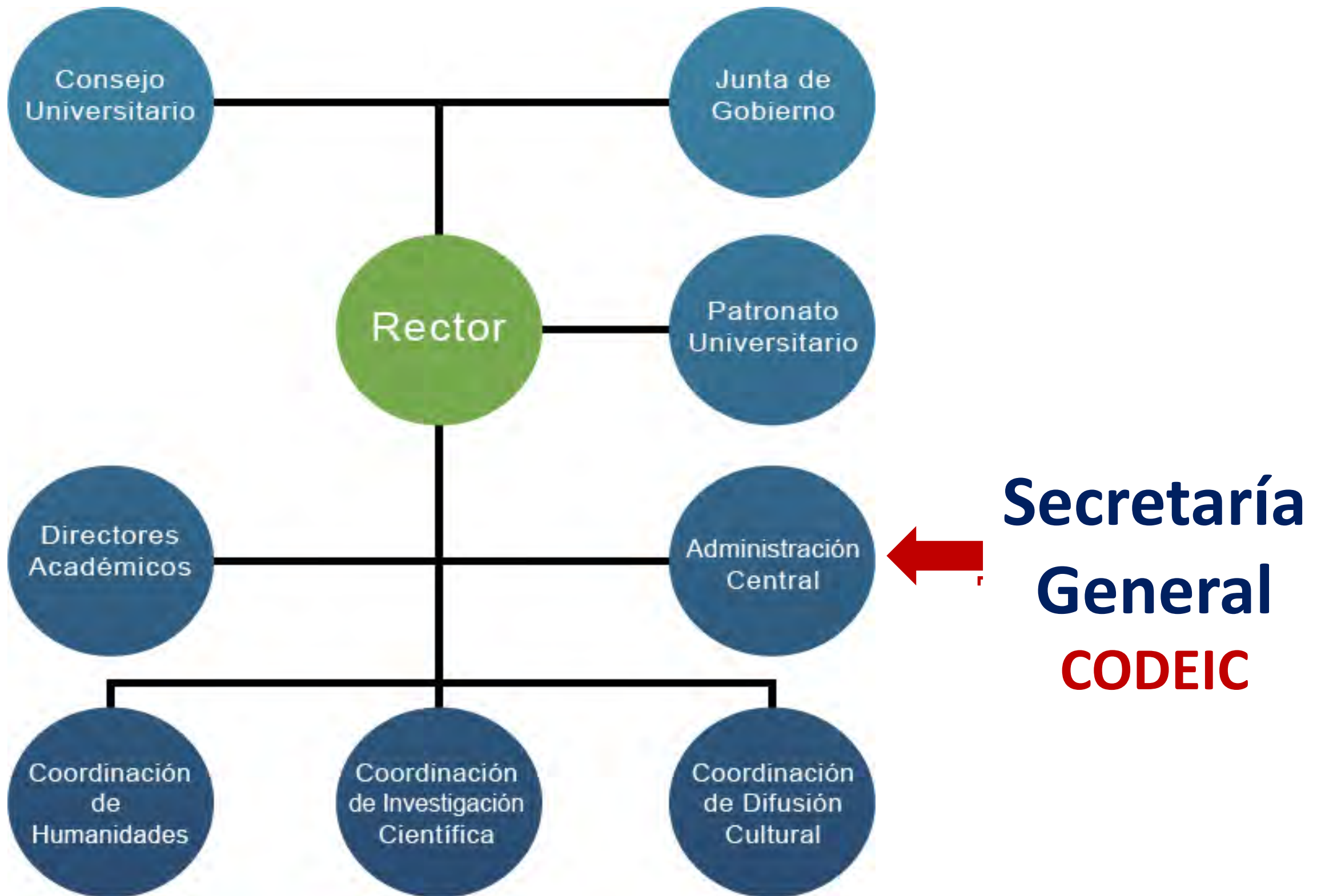
Gaceta **15**



Universidad Nacional Autónoma de México
Dirección General de Evaluación Educativa

QUINTO.- La Dirección General de Evaluación Educativa forma parte de la Coordinación de Desarrollo Educativo e Innovación Curricular y conserva sus funciones.

Organigrama UNAM



Misión

“Participar con la UNAM en el desarrollo, implementación y evaluación de innovaciones educativas, estrategias curriculares, métodos de enseñanza y evaluación, basados en la mejor evidencia disponible, para impulsar el desarrollo profesional docente y el aprendizaje complejo de los estudiantes”.

DDEIC ↔ DEE



Impulso a la docencia

La UNAM instaló el Consejo de Evaluación Educativa





evaluación
de los
aprendizajes
en el
bachillerato

un compendio
de buenas
prácticas



http://www.codeic.unam.mx/wp-content/uploads/2019/03/evaluacion_de_los_aprendizajes.pdf

Diplomado en evaluación del y para el aprendizaje

120^h



Modalidad semipresencial

Objetivos

- 1 Actualizar el marco de referencia del profesorado de educación media superior y superior sobre la evaluación del y para el aprendizaje en los escenarios educativos del siglo XXI.
- 2 Fortalecer sus competencias metodológicas para diseñar e instrumentar programas, estrategias y acciones válidas, pertinentes y objetivas de evaluación del y para el aprendizaje.
- 3 Debatir tendencias e innovaciones en la evaluación del y para el aprendizaje en los escenarios educativos de educación media superior y superior.

Responsables académicos:

Dr. Melchor Sánchez Mendiola
Dr. Adrián Martínez González

Evaluación educativa del y para el aprendizaje en educación superior



www.coursera.org/learn/evaluacion-educativa



Universidad Nacional
Autónoma de México



Prueb@te

LICENCIATURA **UNAM**



Preregistrarse



Inicia sesión ▶

Hecho en México, Universidad Nacional Autónoma de México (UNAM), todos los derechos reservados 2017. Esta página puede ser reproducida con fines no lucrativos, siempre y cuando no se mutile, se cite la fuente completa y su dirección electrónica. De otra forma, requiere permiso previo por escrito de la institución.



Sitio Web administrado por: Coordinación de Desarrollo Educativo e Innovación Curricular (CODEIC) Créditos

www.pruebate.unam.mx

Evaluación diagnóstica

The screenshot displays the website interface for the diagnostic evaluation. At the top, there is a blue header with the UNAM logo and name on the left, a search bar in the center, and the CODEIC logo on the right. The main content area features the title 'Diagnóstico' and the subtitle 'RESULTADOS DEL DIAGNÓSTICO DE CONOCIMIENTOS DE LA LICENCIATURA'. Below this, there are two preview cards. The left card shows the UNAM logo, the CODEIC logo, and the text 'Universidad Nacional Autónoma de México, Secretaría General, Coordinación de Desarrollo Educativo e Innovación Curricular'. A blue button on this card reads 'Manejo del anexo de tablas y gráficos dinámicos del examen diagnóstico de licenciatura de la generación 2018'. The right card is orange and features the UNAM logo, the text 'UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO', and 'Exámenes para el diagnóstico de conocimientos'. At the bottom of the page, there is a date 'Mayo 2018' and the text 'Resultados de los alumnos'.

<https://www.codeic.unam.mx/index.php/diagnostico/>

La evaluación del aprendizaje de los estudiantes: ¿es realmente tan complicada?

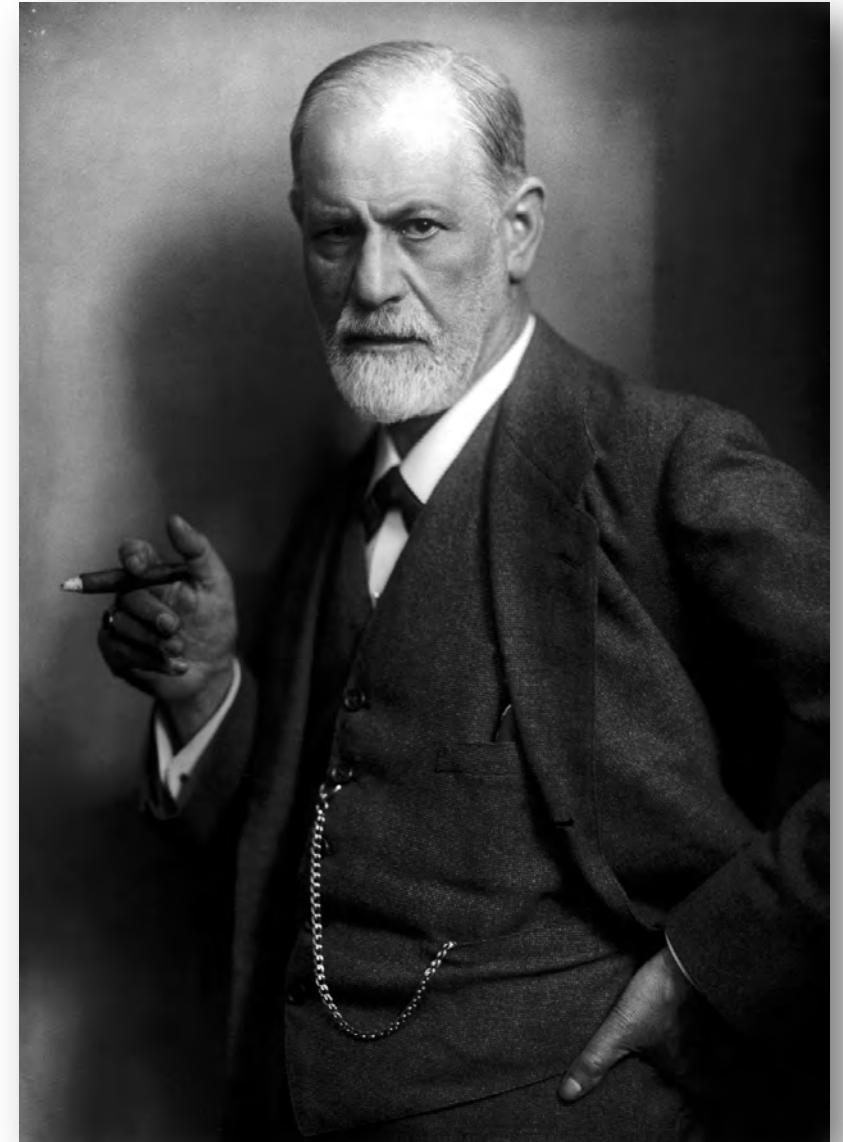


Melchor Sánchez Mendiola

Panorama de la sesión...

- ✓ **Selección para la educación superior: retos evaluativos**
- ✓ Examen de ingreso a la licenciatura de la UNAM: Evidencia de validez
- ✓ Algunas reflexiones finales

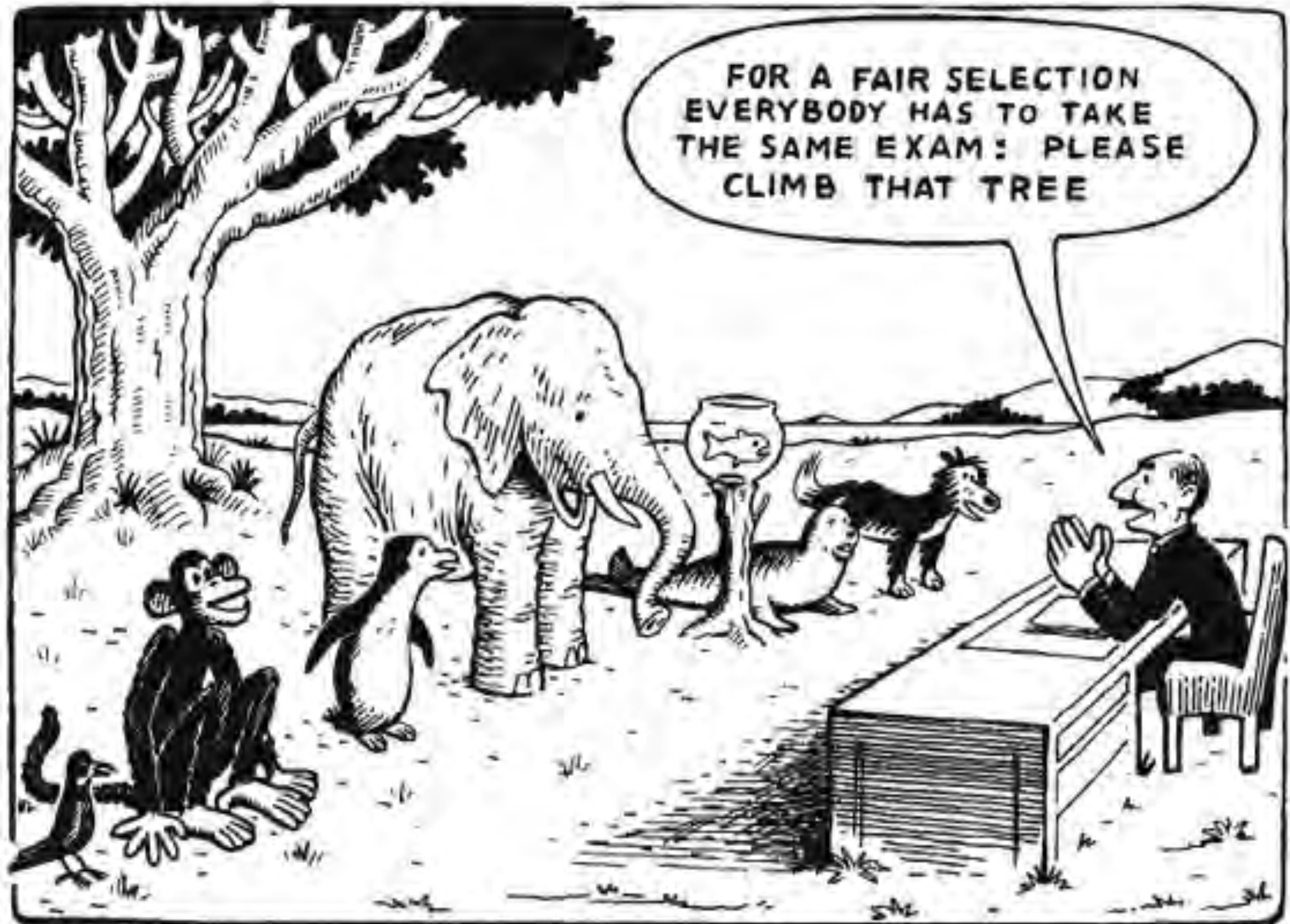
“Parecería que el psicoanálisis fuera una de esas profesiones “imposibles” en las que puedes estar seguro de antemano de lograr resultados insatisfactorios. Las otras dos, que se han conocido por más tiempo, son la **educación** y el **gobierno**”.



Sigmund Freud, 1937

“Analysis Terminable and Interminable”

FOR A FAIR SELECTION
EVERYBODY HAS TO TAKE
THE SAME EXAM: PLEASE
CLIMB THAT TREE



Consenso de Ottawa 2018: Selección y reclutamiento para las profesiones de la salud

Métodos

- 1) Logro académico
- 2) Exámenes de aptitudes
- 3) Declaraciones personales y CV
- 4) Referencias y cartas de recomendación
- 5) Exámenes de juicio situacional (SJT)
- 6) Evaluación de la personalidad
- 7) Entrevistas; MMI (mini-entrevista múltiple)
- 8) Centros de selección (muestras de trabajo y simulación)

Patterson F, et al. Medical Teacher 2018; 40(11):1091-1101.

Sólo 8.6% de aspirantes a la UNAM se quedaron

- La máxima casa de estudios dio a conocer los resultados de su primer examen de admisión a cualquiera de sus 118 carreras



Primer concurso de ingreso al nivel licenciatura para el ciclo 2017. Acudieron 144 mil 61 estudiantes, sólo el 8.6 % alcanzó un lugar.

<http://www.eluniversal.com.mx/articulo/nacion/sociedad/2017/03/27/solo-86-de-aspirantes-la-unam-se-que-daron>

**“ El poder de
examinar es el
poder de
destruir...”**



**Abraham Flexner. Medical education
in the United States and Canada.**

1910. p. 169

www.carnegiefoundation.org

Indian parents scale school wall to help students cheat on exams



<https://www.cbsnews.com/news/indian-parents-scale-school-wall-to-help-students-cheat-on-exams/>



Investigación en
Educación Médica

<http://riem.facmed.unam.mx>



ARTÍCULO DE REVISIÓN

Exámenes de alto impacto: implicaciones educativas

Melchor Sánchez-Mendiola^{a,*} y Laura Delgado-Maldonado^b

Inv Ed Med. 2017;6(21):52–62

<https://www.sciencedirect.com/science/article/pii/S2007505716300850>

Exámenes de Altas Consecuencias y sus efectos: Calles de dos sentidos



Cizek, G.J. (2001). More unintended consequences of high-stakes testing. *Educ Meas* 20:19-27

Consecuencias de un Examen

Intencionales No intencionales

Positivas

I - P

NI - P

Negativas

I - N

NI - N

**Brennan RL, Ed. Educational Measurement.
National Council on Measurement in Education
4th Ed. 2006**

“Los desarrolladores del examen son los candidatos obvios para validar las afirmaciones que hacen sobre la interpretación de los resultados...”

**Brennan RL, Ed. Educational Measurement.
National Council on Measurement in Education
4th Ed. 2006**

Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference

- ✓ Validez
- ✓ Confiabilidad
- ✓ Equivalencia
- ✓ Factibilidad
- ✓ Efecto educativo
- ✓ Efecto catalítico
- ✓ Aceptabilidad

2011; 33: 206–214

**MEDICAL
TEACHER**

Validez

“Validez es un juicio evaluativo e integrado, del grado al que la evidencia empírica y razonamientos teóricos apoyan lo adecuado y apropiado de las inferencias y acciones basadas en puntuaciones de exámenes y otras formas de evaluación”



Samuel
Messick

Messick, 1989

Validez

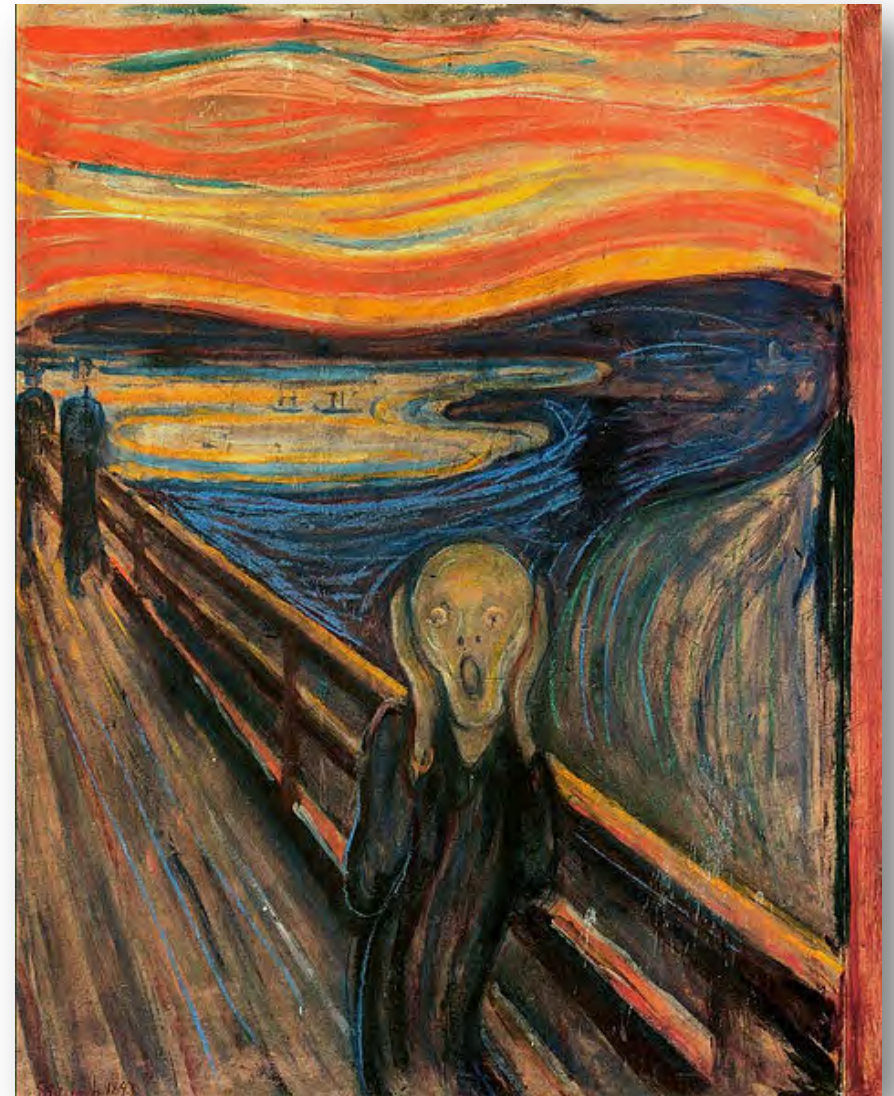
“Validar una interpretación propuesta o un uso de los puntajes de exámenes es evaluar las afirmaciones basadas en esas puntuaciones. La mezcla específica de evidencia necesaria para la validación depende de las inferencias que se hagan y los supuestos que se utilicen”

Michael Kane, 2006



INFERENCIA, ¡no instrumento!

- “El documento sólido”
Usamos el documento
“El documento”
El instrumento
do ya que
un... ach
arco



Principios de Validez

- **Toda la validez es validez de constructo**
- Investigación de validez: más (o menos) evidencia para (o contra) usos específicos de los resultados de evaluación
- A mayor consecuencia de un examen, se requiere más evidencia de validez

Cinco Fuentes de Evidencia

- 1) Contenido del examen
- 2) Proceso de respuesta
- 3) Estructura interna
- 4) Relación con otras variables
- 5) Consecuencias del examen

AERA, APA, NCME. Standards for Educational and Psychological Testing, 2014

Validez

“La responsabilidad final del uso e interpretación apropiada de los exámenes descansa predominantemente en el usuario de los resultados”.

AERA, APA & NCME. Standards for Educational and Psychological Testing. Washington, DC. 2014

Panorama de la sesión...

- ✓ Selección para la educación superior: retos evaluativos
- ✓ **Examen de ingreso a la licenciatura de la UNAM: Evidencia de validez**
- ✓ Algunas reflexiones finales

Reglamento General de Inscripciones UNAM

Art. 2. Para ingresar a la Universidad es indispensable:

- Ser aceptado mediante **concurso de selección**, que comprenderá una **prueba escrita** y que deberá realizarse dentro de los periodos que al efecto se señalen.

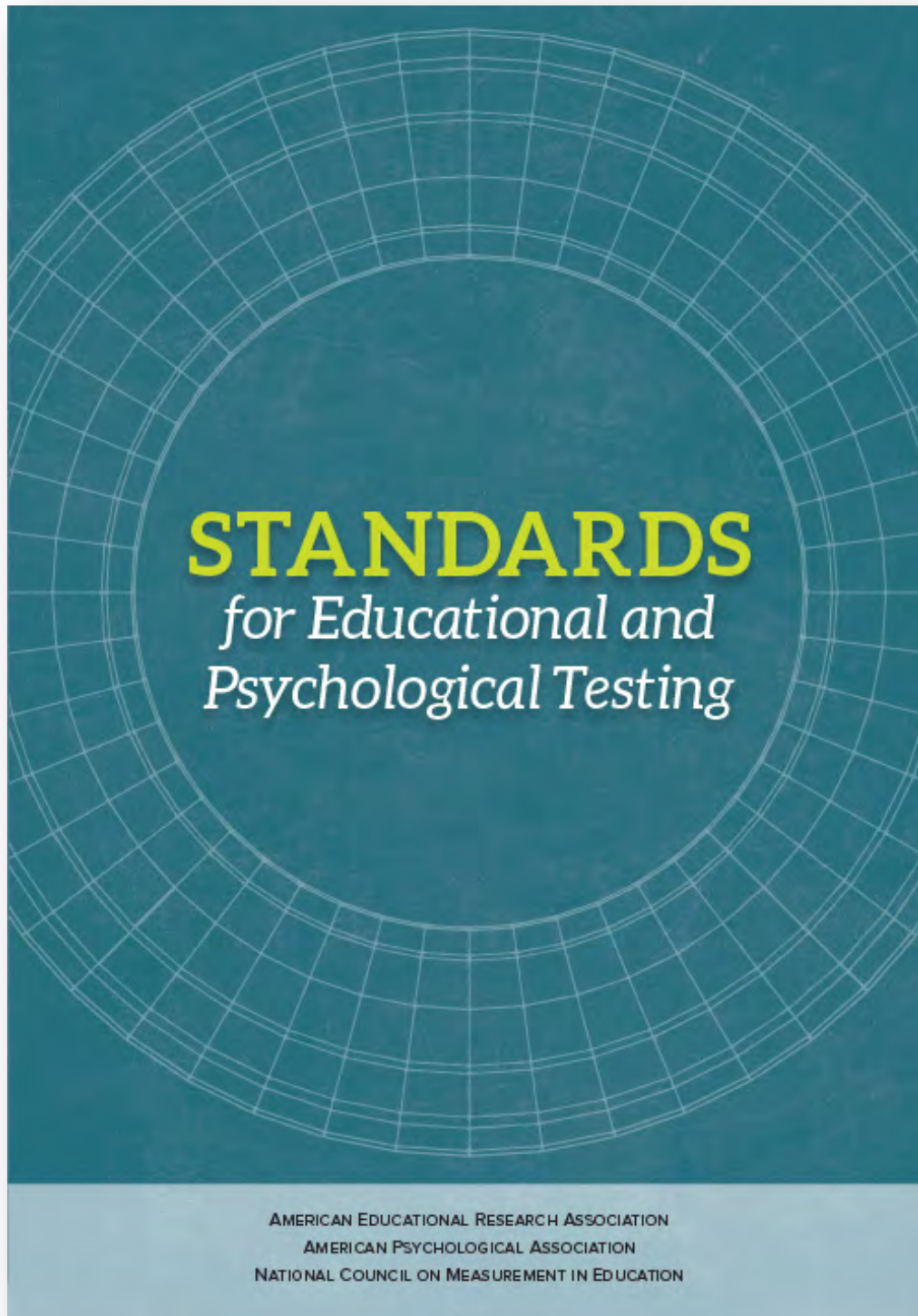
<https://www.dgae-siae.unam.mx/acerca/normatividad.html#leg-3>

2nd Edition

Handbook of Test Development

Edited by
Suzanne Lane, Mark R. Raymond,
and Thomas M. Haladyna





AMERICAN
PSYCHOLOGICAL
ASSOCIATION



www.aera.net/Publications/tabid/10067/Default.aspx

Metodología del diseño de los exámenes de ingreso



Contenido del examen

- Fundamentado en los planes de estudio de la educación media superior.
- Comisiones de profesores del bachillerato de la UNAM revisaron los temarios con un perfil de referencia y determinaron los contenidos a evaluar.
- Tabla de especificaciones.
- Entrenamiento de elaboradores de reactivos.

Proceso de respuesta

- Familiaridad del estudiante con el formato POM.
- Cuatro opciones de respuesta, una correcta.
- Revisado por CODEIC para verificar congruencia, relación con el resultado de aprendizaje, estructura gramatical.
- Validación de la clave de respuestas.
- Control de calidad del reporte de resultados.

Sistema Integral de Gestión de Exámenes (SIGE)

Revisión interna

Reactivo de repositorio Propuesta del profesor **Edición del reactivo**

Enunciado Despliegue de respuestas Opción A) Opción B) Opción C) Opción D) Instrucción Lectura Imágenes repositorio

Galería Repositorio

Formato de texto: Arial, Tamaño del texto

Campo para la base del reactivo

Actualizar vista previa

Guardar cambios

Área de
captura
del
reactivo

Guardar datos de revisión interna Regresar al listado

Criterio	Válido	Relación de comentarios	Agregar comentario
Respuesta correcta A)			
Justifique por qué sin lugar a duda esta es la única respuesta correcta.	SI		
La opción correcta está redactada con claridad.	SI		

Espacio para observaciones por parte del personal de CODEIC

Sistema Integral de Gestión de Exámenes (SIGE) de la CODEIC

- Cuando el reactivo cumple con los criterios psicométricos, es revisado por tres profesores expertos en el tema para determinar si el contenido que evalúa el reactivo es correcto.

Estructura interna

- Confiabilidad
- Análisis de reactivos
 - ✓ Dificultad y discriminación de los ítems
 - ✓ Análisis de las opciones (DNF)
- Error estándar de medición (SEM)
- Análisis factorial
- Funcionamiento diferencial de ítems (DIF)

Relación con otras variables

- El examen de ingreso se asocia con los resultados de los exámenes de diagnóstico de conocimientos.
- La correlación entre los exámenes de ingreso y Dx Conocimientos es $r = 0.64$ ($p < 0.01$)

Relación con otras variables

Correlaciones por área de conocimiento entre el examen de ingreso y el examen diagnóstico de conocimientos:

ÁREA	CORRELACIÓN
FMI	0.62*
CBQS	0.60*
CS	0.65*
H y A	0.77*

*p < 0.01

Consecuencias del examen

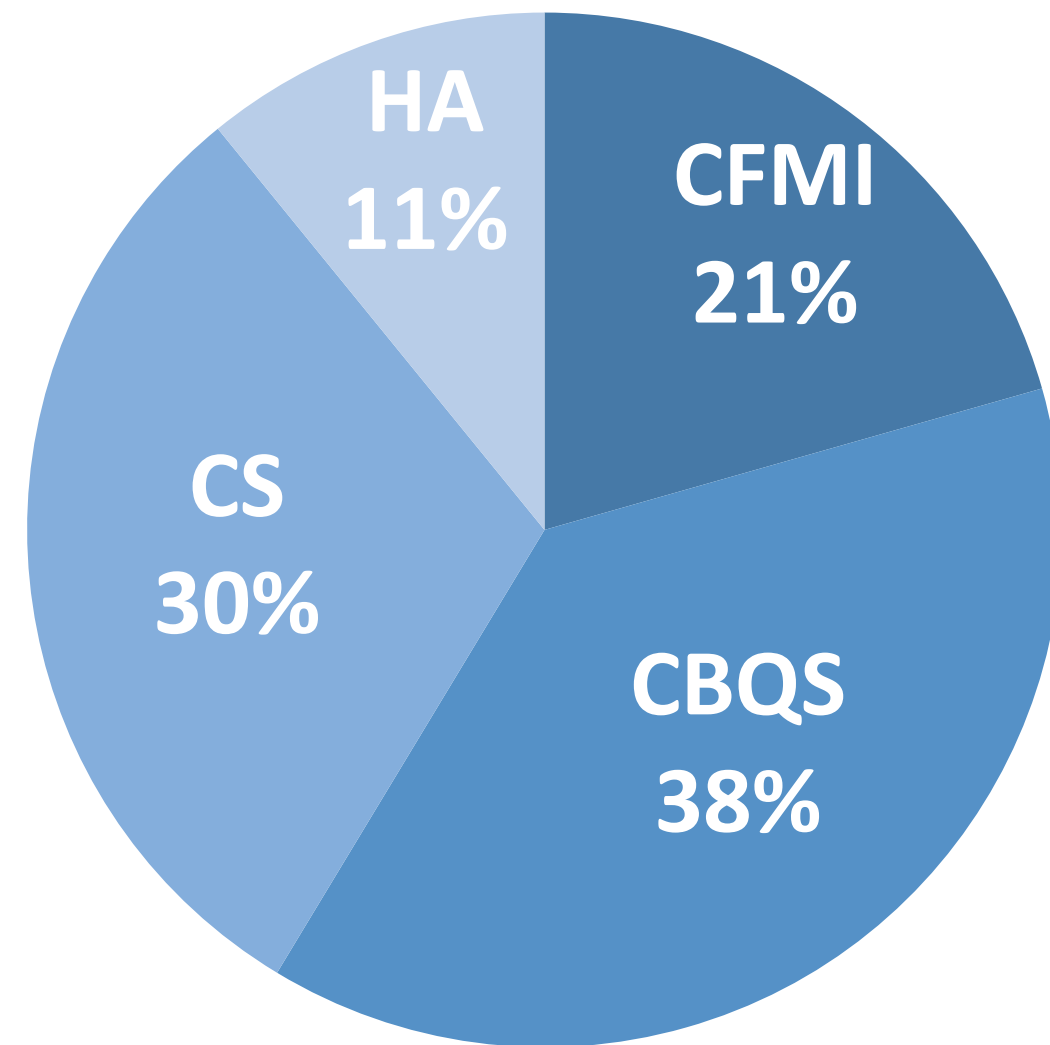
- Evidencia de efectos de exámenes en estudiantes, escuelas y Sociedad (publicaciones)
- Efectos de ingresar o no:
 - ✓ Costos económicos y emocionales
 - ✓ Costos sociales de falsos positivos y falsos negativos
- No se realiza una determinación criterial de puntuación de pase.

Resultados del examen de ingreso febrero 2019

- 1) Distribución de aspirantes.
- 2) Análisis psicométrico.
- 3) Resultados de los aspirantes.
- 4) Análisis de comportamiento diferencial de los ítems (DIF).

Aspirantes por área

Área	N	%
CFMI	30,561	20.6
CBQS	56,474	38.1
CS	45,229	30.5
HyA	16,143	10.9
Total	148,407	100.0



CFMI=Ciencias Físico Matemáticas e Ingenierías; CBQS=Ciencias Biológicas, Químicas y de la Salud; CS=Ciencias Sociales; HyA=Humanidades y Artes.

Aspirantes por día y turno

Día	Turno		Total
	Matutino	Vespertino	
1	37,235 (25.1)	36,908 (24.9)	74,143 (50.0)
2	36,356 (24.5)	37,908 (25.5)	74,264 (50.0)
Total	73,591 (49.6)	74,816 (50.4)	148,407 (100.0)

Las cifras entre paréntesis representan el porcentaje con respecto al total de aspirantes del concurso (148,407).

Aspirantes por sexo

	Área				Total
	CFMI	CBQS	CS	HyA	
Hombres	21,486	18,082	20,723	5,638	65,929
%	32.6%	27.4%	31.4%	8.6%	100.0%
% en el área	70.3%	32.0%	45.8%	34.9%	44.4%
Mujeres	9,075	38,392	24,506	10,505	82,478
%	11.0%	46.5%	29.7%	12.7%	100.0%
% en el área	29.7%	68.0%	54.2%	65.1%	55.6%
Total	30,561	56,474	45,229	16,143	148,407
%	20.6%	38.1%	30.5%	10.9%	100.0%

Estructura del Examen

Materia	CFMI	CBQS	CS	H y A
Matemáticas	26	24	24	22
Física	16	12	10	10
Química	10	13	10	10
Biología	10	13	10	10
H. universal	10	10	14	10
H. de México	10	10	14	10
Literatura	10	10	10	10
Geografía	10	10	10	10
Español	18	18	18	18
Filosofía	-	-	-	10
Total	120	120	120	120

Descriptivos por área de conocimiento

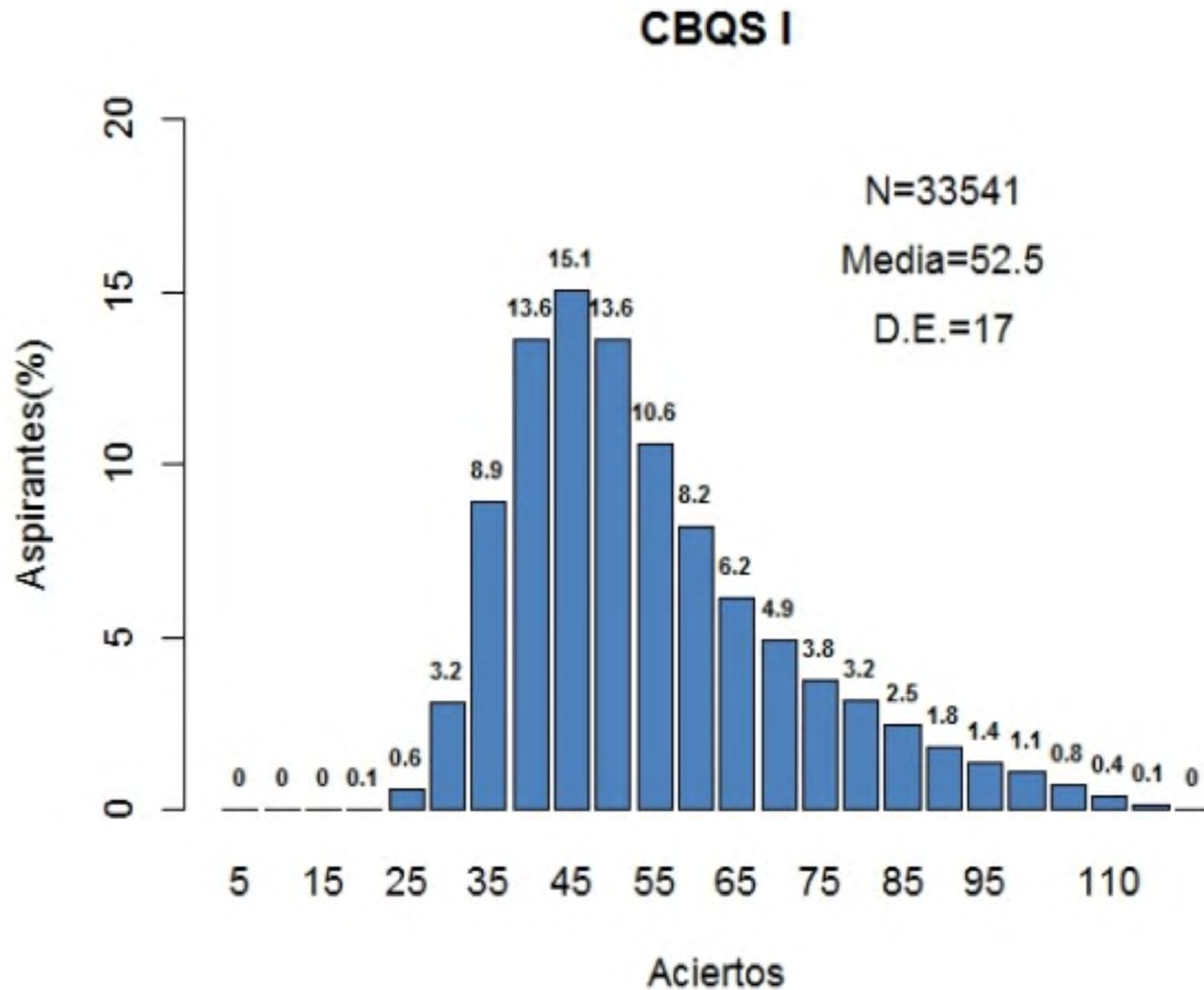
Área	Aspirantes		# de aciertos			
	N	%	Promedio	D.E.	Min	Max
CFMI	30,561	20.6	56.5	18.2	16	117
CBQS	56,474	38.1	52.7	17.0	1	117
CS	45,229	30.5	54.7	16.4	0	115
HyA	16,143	10.9	54.9	16.4	9	114
Total	148,407	100.0	54.3	17.0	0	117

CFMI=Ciencias Físico Matemáticas e Ingenierías; CBQS=Ciencias Biológicas, Químicas y de la Salud; CS=Ciencias Sociales; HyA=Humanidades y Artes.

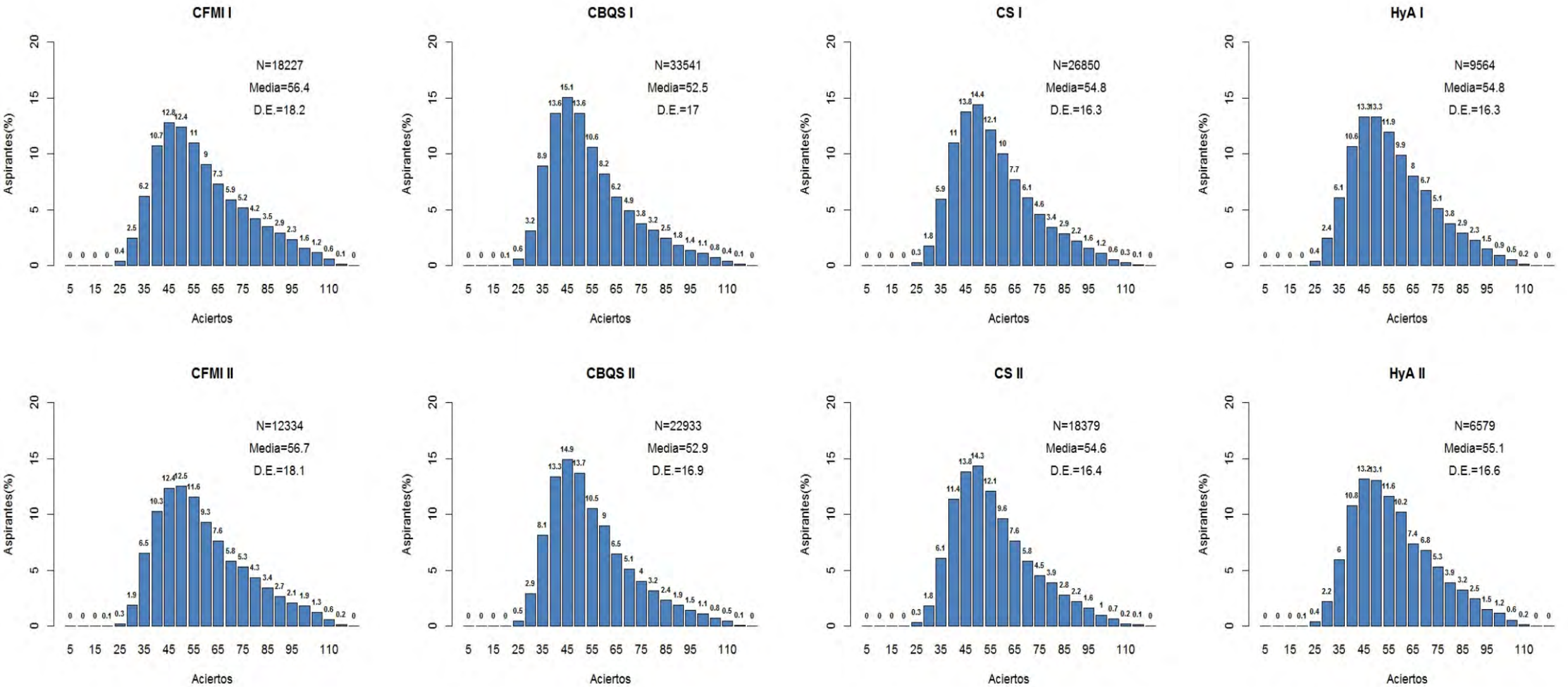
Análisis psicométrico Examen de Ingreso (TMC)

	CFMI		CBQS		CS		HyA	
	I	II	I	II	I	II	I	II
N	18,227	12,334	33,541	22,933	26,850	18,379	9,564	6,579
Promedio aciertos	55.4	55.7	52.5	52.9	54.8	54.6	54.8	55.1
Desv. Estándar	18.2	18.1	17.0	16.9	16.3	16.4	16.3	16.6
Mediana	52	52	49	49	52	51	52	52
Dificultad media	0.466	0.468	0.437	0.441	0.456	0.455	0.456	0.460
Correlación punto biserial media	0.298	0.296	0.273	0.271	0.260	0.263	0.258	0.264
Alfa de Cronbach	0.927	0.926	0.914	0.913	0.907	0.908	0.906	0.910

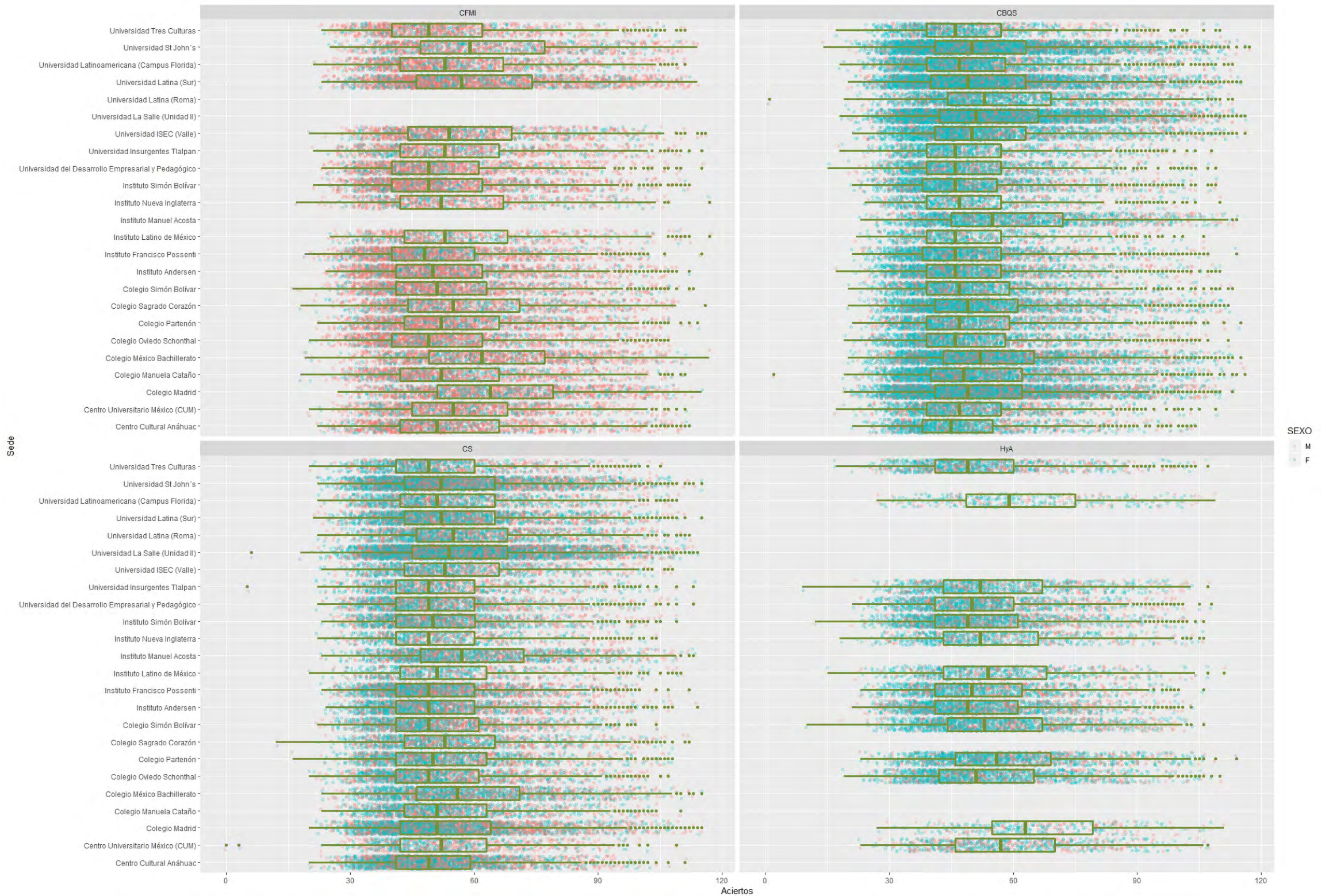
Distribución de aciertos por área y examen



Distribución de aciertos por área y examen

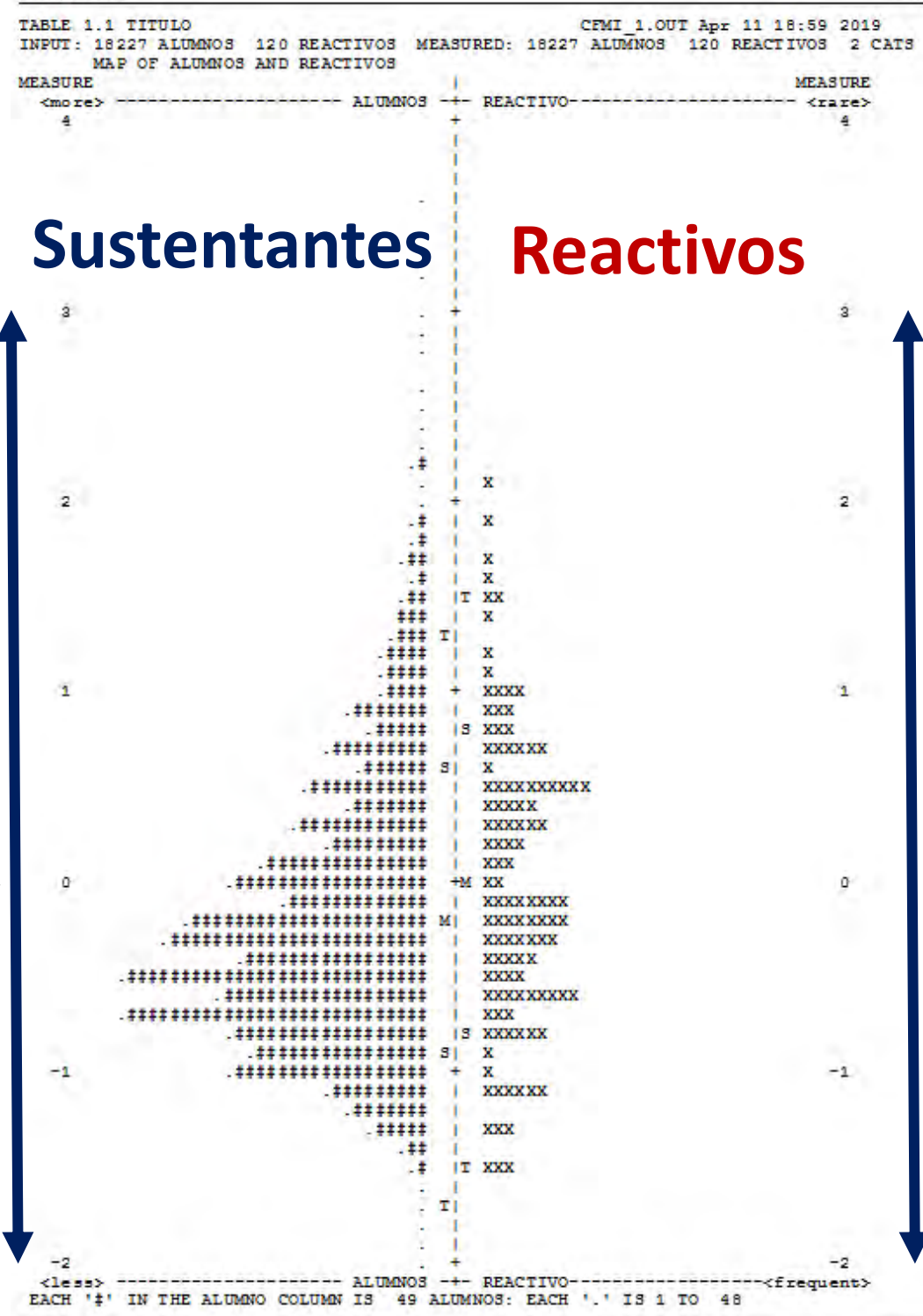


Puntajes de los aspirantes por sede



Mapa de dificultad (Teoría de respuesta al ítem)

CFMI



Sustentantes

Reactivos

Mayor habilidad

Mayor dificultad

Media

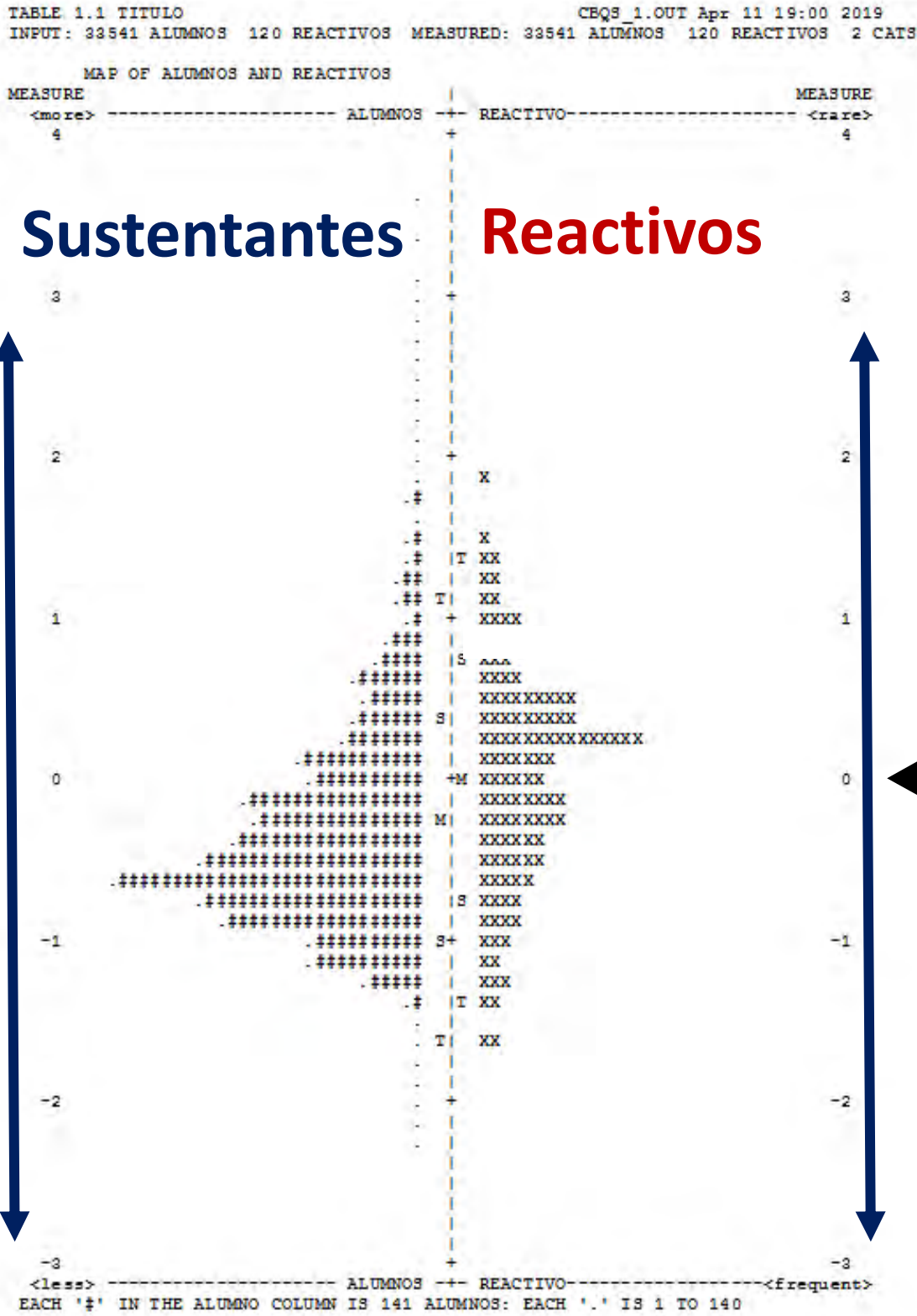
Media

Menor habilidad

Menor dificultad

Mapa de dificultad (Teoría de respuesta al ítem)

CBQS



Mayor habilidad

Mayor dificultad

Media →

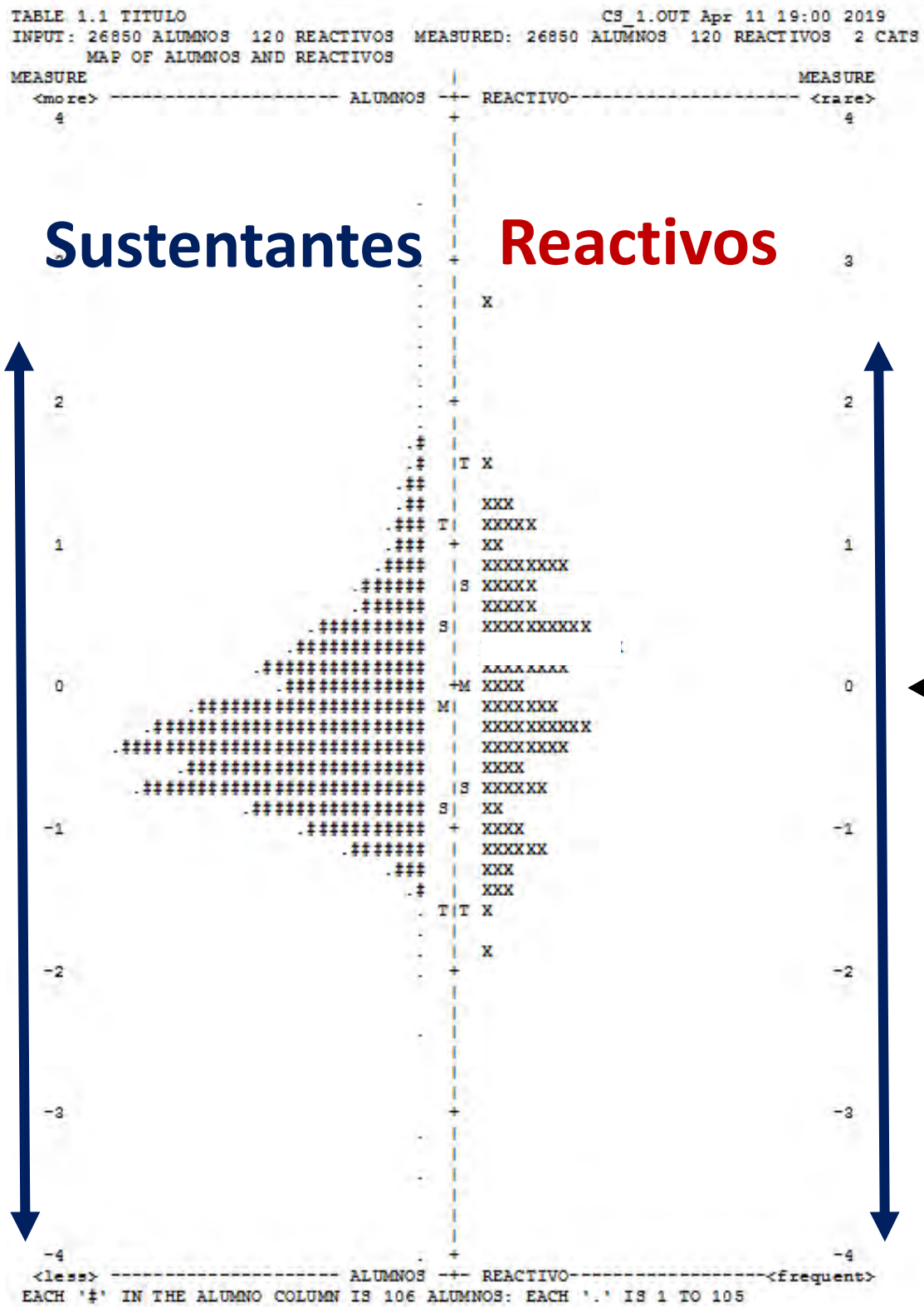
← Media

Menor habilidad

Menor dificultad

Mapa de dificultad (Teoría de respuesta al ítem)

CS



Mayor habilidad

Mayor dificultad

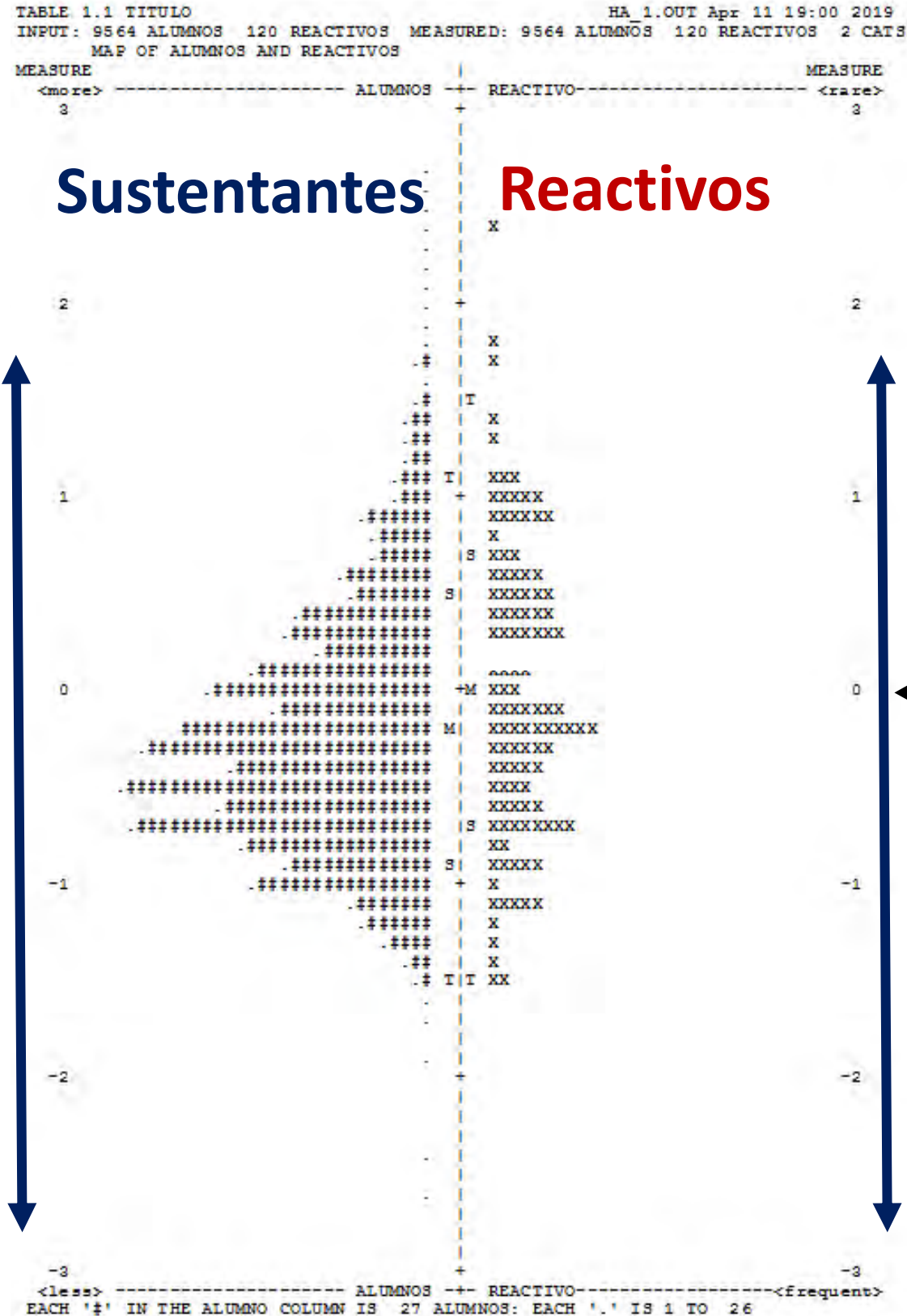
Media →

← Media

Menor habilidad

Menor dificultad

Mapa de dificultad (Teoría de respuesta al ítem)



HA

Mayor habilidad

Media

Menor habilidad

Mayor dificultad

Media

Menor dificultad

Descriptivos por sexo

Sexo	Aspirantes		# de aciertos	
	N	%	Promedio	D.E.
Hombres	65,929	44.42%	56.4	17.8
CFMI	21,486	32.59%	56.9	18.5
CBQS	18,082	27.43%	55.3	18.0
CS	20,723	31.43%	56.6	16.9
HA	5,638	8.55%	57.6	16.9
Mujeres	82,478	55.58%	52.7	16.3
CFMI	9,075	11.00%	55.8	17.3
CBQS	38,392	46.55%	51.4	16.3
CS	24,506	29.71%	53.1	15.7
HA	10,505	12.74%	53.5	16.0
Total	148,407	100.00%	54.3	17.0

**Análisis de funcionamiento
diferencial (DIF) de los
reactivos del examen de
ingreso a las licenciaturas de la
UNAM por sexo del aspirante**

DIF

- Un reactivo presenta DIF (*Differential Item Functioning*) cuando los examinados de un mismo nivel de habilidad, pero provenientes de diferentes grupos, tienen una probabilidad distinta de contestarlo correctamente.

DIF

- Se empleó la técnica de TRI basada en el modelo de Rasch, donde los grupos de interés fueron las mujeres y los hombres.
- Si los contrastes en los niveles de dificultad de un reactivo entre los grupos de interés no superaba los 0.43 lógitos, DIF = sin importancia; > 0.64 lógitos, DIF = moderado-alto; encontrarse entre esos valores DIF = leve-moderado. (*Educational Testing Service*)

Reactivos con DIF

Filosofía

FIL02383
FIL04043
FIL09404
FIL08673
FIL04863
FIL07335
FIL09395
FIL08859
FIL04026
FIL07646
FIL10029
FIL10030
FIL08973
FIL06475
FIL10031
FIL10032

Hombres



Moderado-Alto
Leve-Moderado

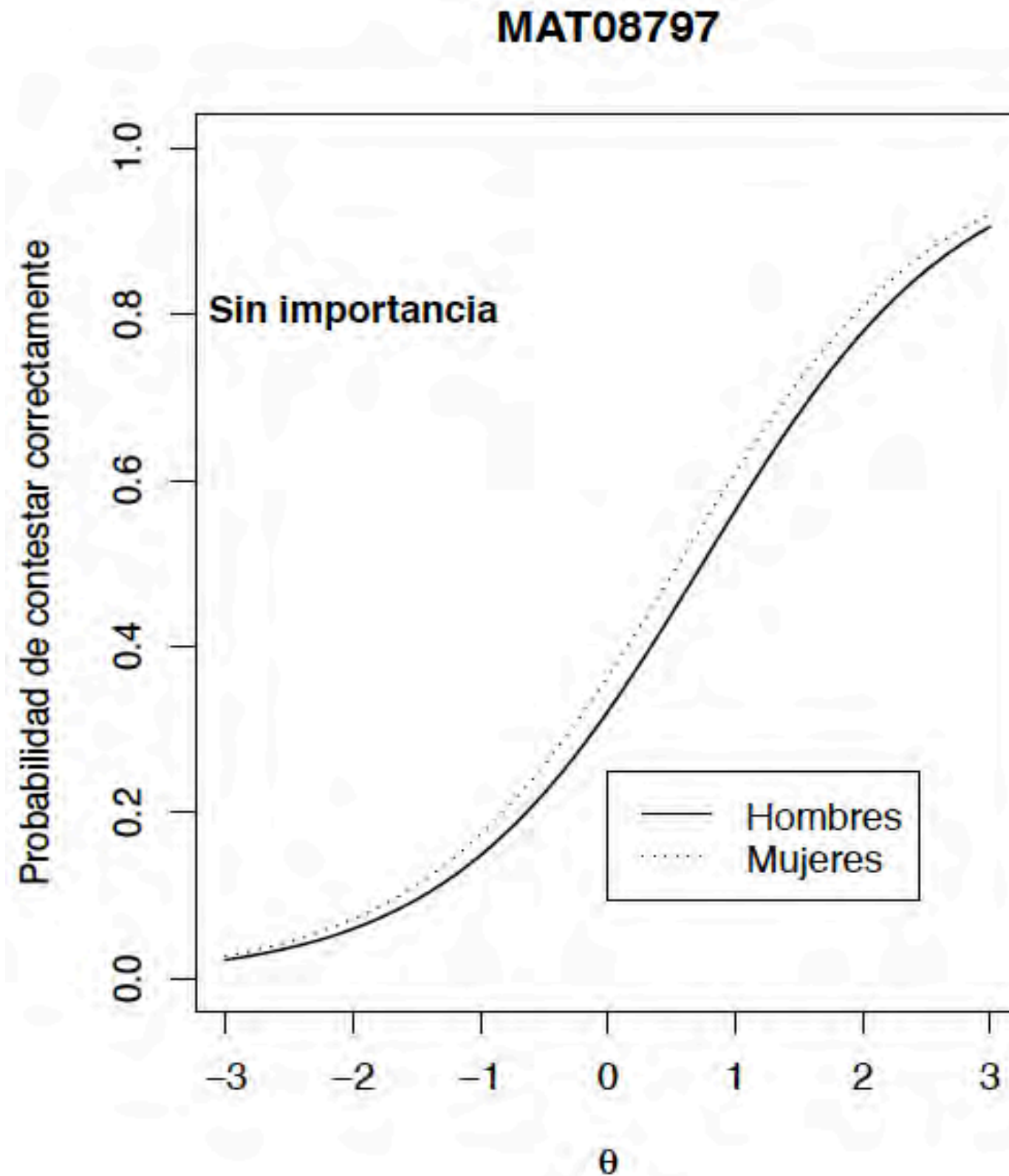
Sin importancia

Mujeres

FIL02383
FIL04043
FIL09404
FIL08673
FIL04863
FIL07335
FIL09395
FIL08859
FIL04026
FIL07646
FIL10029
FIL10030
FIL08973
FIL06475
FIL10031
FIL10032

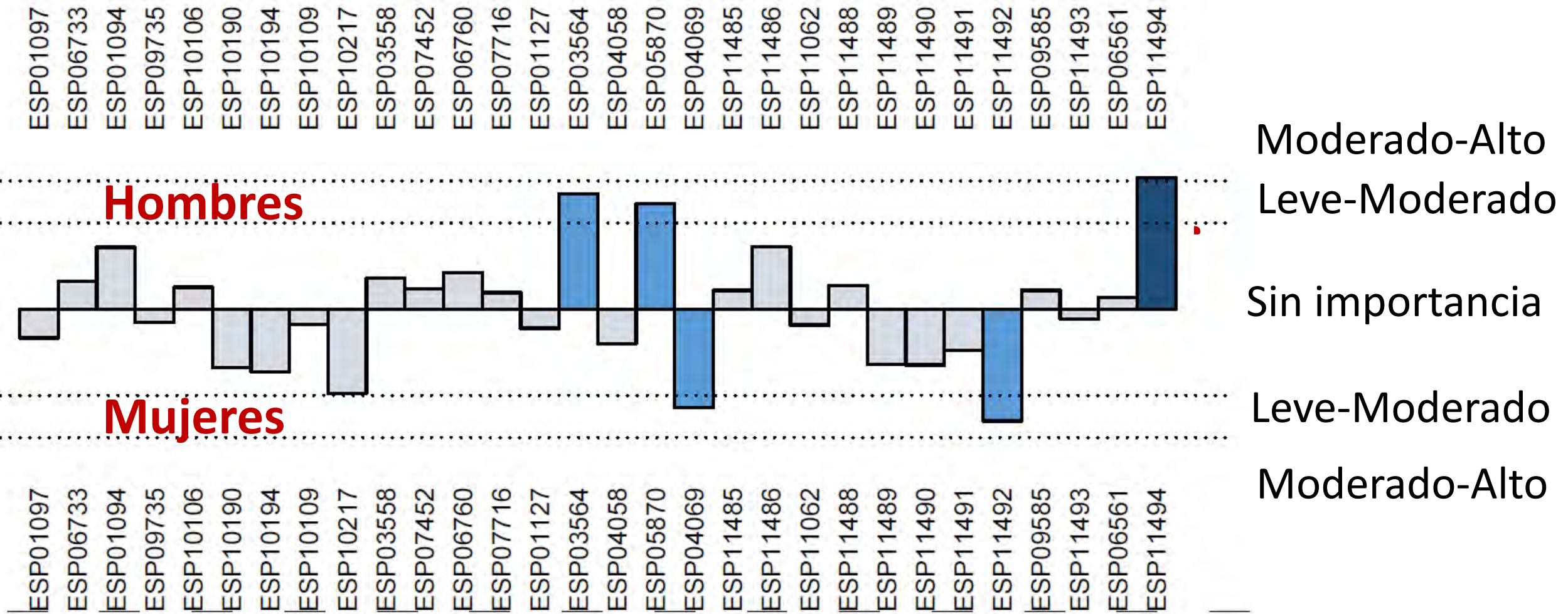
Leve-Moderado
Moderado-Alto

Reactivos con DIF sin importancia



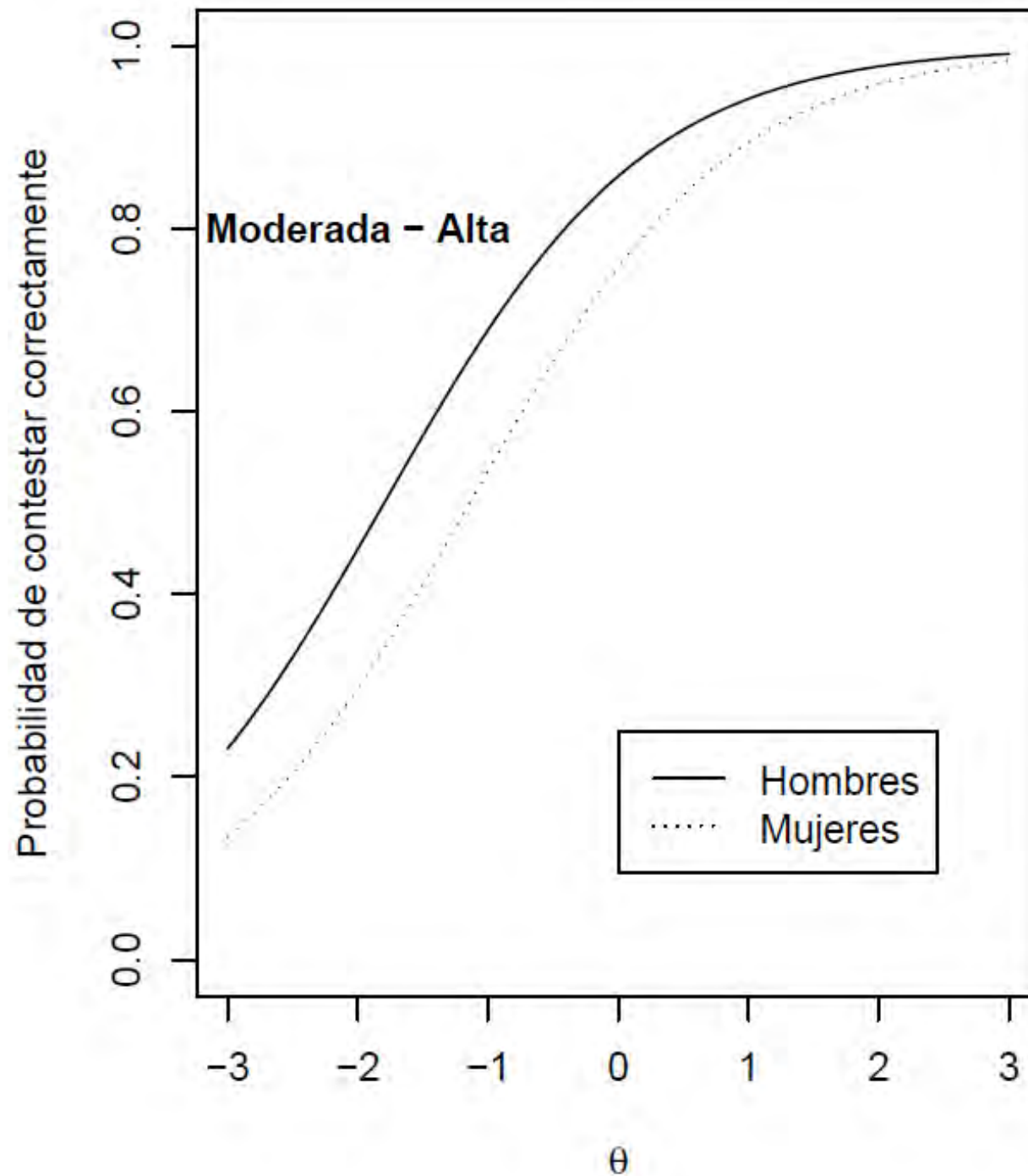
Reactivos con DIF moderado - alto

Español CBQS



Reactivos con DIF moderado - alto

ESP11494



Panorama de la plática...

- ✓ Selección para la educación superior: retos evaluativos
- ✓ Examen de ingreso a la licenciatura de la UNAM: Evidencia de validez
- ✓ **Algunas reflexiones finales**

Recursos empleados

- Para la elaboración del examen participan aprox. 85 profesores de bachillerato UNAM.
- 10 responsables de asignatura de la CODEIC.
- Tres correctores de estilo.
- En el piloteo de pruebas apoyan 25 aplicadores coordinados por CODEIC.
- Personal de sistemas y evaluación educativa.
- Aproximadamente 9 meses de trabajo.

Assessment in the post-psychometric era: Learning to love the subjective and collective

- ✓ S. XX el péndulo de la evaluación se movió en exceso a lo cuantitativo (“psicometría”).
- ✓ Lo “objetivo” es bueno y lo “subjetivo” es malo (sesgado, injusto, no estandarizado).
- ✓ Problemas: exceso de confianza en los números; malinterpretación de resultados; exagerada necesidad de seguridad; altos costos; intervalos largos entre exámenes, hasta el final de cursos.

Hodges B. Med Teach 2013; 35:564-8.

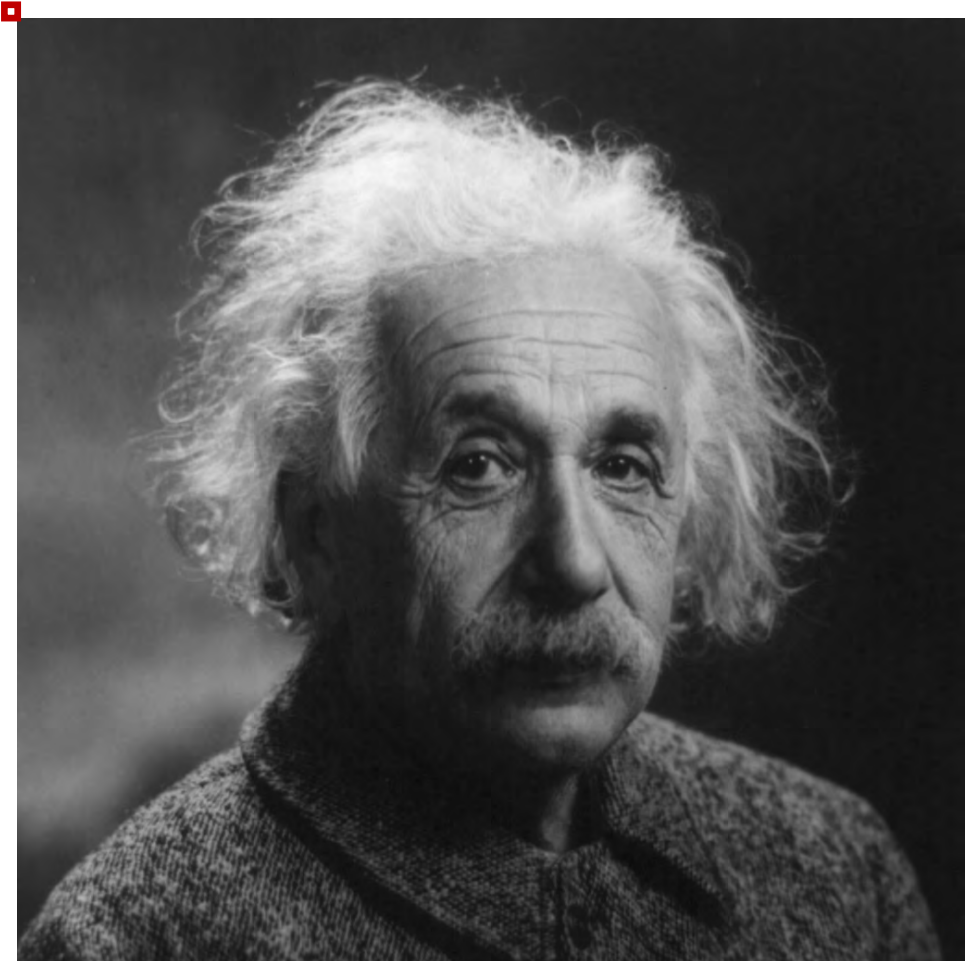
Assessment in the post-psychometric era: Learning to love the subjective and collective

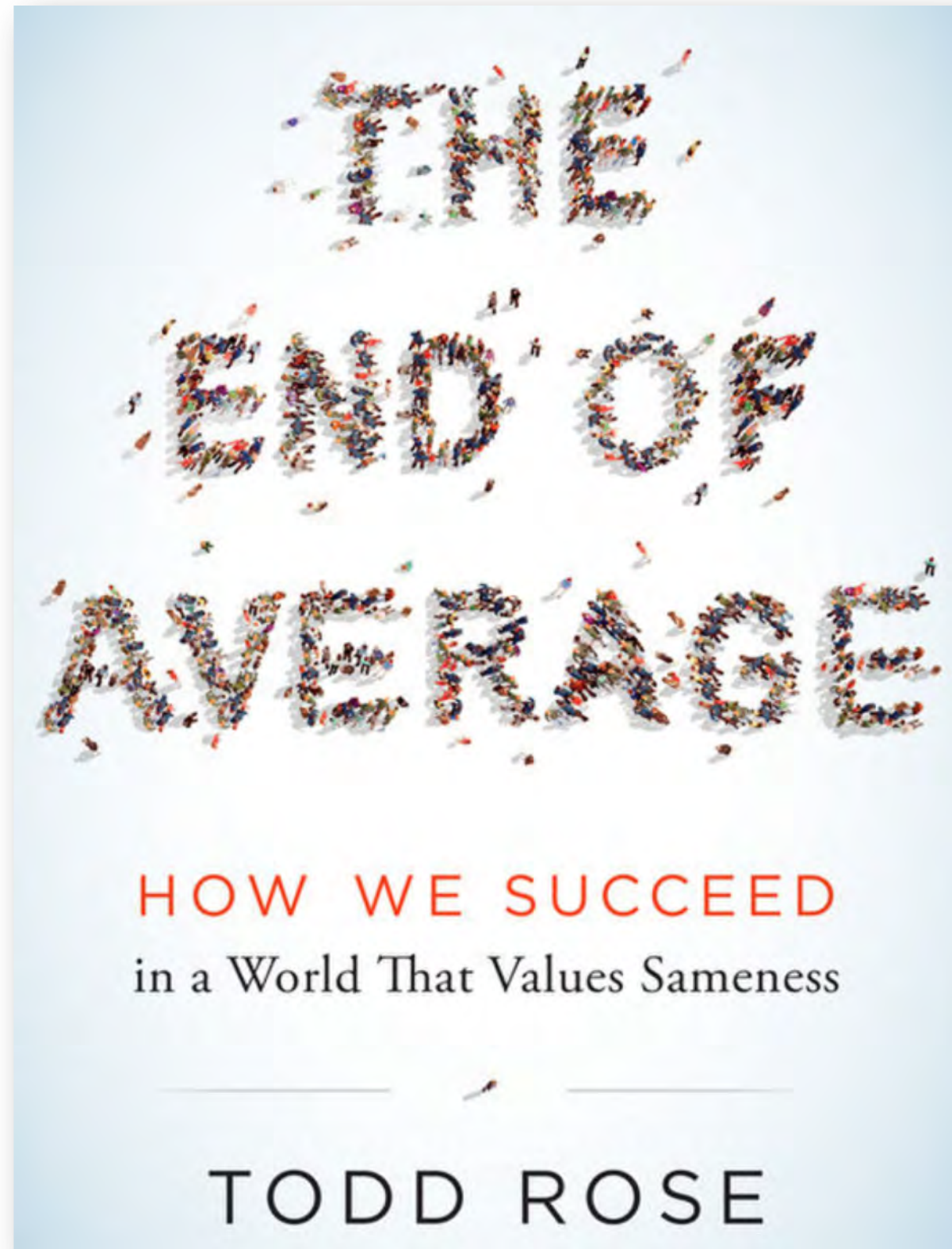
- ✓ Premisa del discurso psicométrico: convertir los fenómenos humanos en números.
- ✓ No es un proceso exacto. Se pierde información.
- ✓ Los números representan algo, pero no son la entidad por sí mismos (CI 115; 7 en Español).
- ✓ Mediciones oscurecidas por ruido estadístico, factores externos e internos al estudiante.

Hodges B. Med Teach 2013; 35:564-8.

**“No todo lo que
puede ser
contado cuenta,
y no todo lo que
cuenta puede ser
contado”**

Albert Einstein





<https://www.harpercollins.com/9780062358363/the-end-of-average>



**“Todo lo que
merece ser hecho,
merece también
ser bien hecho”**

Andre Maurois

¿RECUPERAR LA ESPERANZA?

*La investigación educativa entre pasado y futuro**

PABLO LATAPÍ SARRE

Por esto respondo con un rotundo “sí” a la pregunta de si es posible hoy recuperar la esperanza.

Cumplamos nuestra misión de investigadores comprometidos, con sentido de urgencia. Hay un verso del poeta Jaime Sabines que me viene continuamente a la memoria y por ello me gusta repetirlo.

Dice así: **“La eternidad se nos acaba”**.

RMIE, ENERO-MARZO 2008, VOL. 13, NÚM. 36, PP. 283-297

La eternidad se vuelve finita y se consume y extingue en cada niño que se queda sin escuela, en cada generación perdida, en el desperdicio irreversible del tiempo, recurso no renovable; el tiempo, que es el principal activo de las personas y las sociedades. El tiempo de México, y también el de cada uno de nosotros, se agotan. Vivamos intensamente: que nuestra esperanza, reconstruida en este presente azaroso, crezca ... y se desborde hasta alcanzar el fin de los tiempos.

RMIE, ENERO-MARZO 2008, VOL. 13, NÚM. 36, PP. 283-297