

Procedimientos estadísticos para la evaluación diagnóstica en educación usando R

Miguel A. Sorrel

Universidad Autónoma de Madrid

Study of statistical procedures for diagnostic evaluation in educational contexts

#SmallSampleDiagnosticTesting



Miguel A. Sorrel
Faculty of Psychology
Universidad Autónoma de Madrid
miguel.sorrel@uam.es



Psychometric Models
& Applications Lab.



Universidad Autónoma
de Madrid



Comunidad
de Madrid

SI3/PJI/2021-00258



Pablo
Nájera



Francisco J.
Abad



Rodrigo S.
Kreitchmann



Jimmy
de la Torre

Índice

1. Psicometría
2. Medición diagnóstica
3. Paquetes de R
 - `cdmTools`
 - `cdcatR`
4. Aplicaciones Shiny
 - `FoCo`
 - `cdcatS`
5. Ilustración: Un análisis de inicio a fin
6. Notas finales

Psicometría

Medir

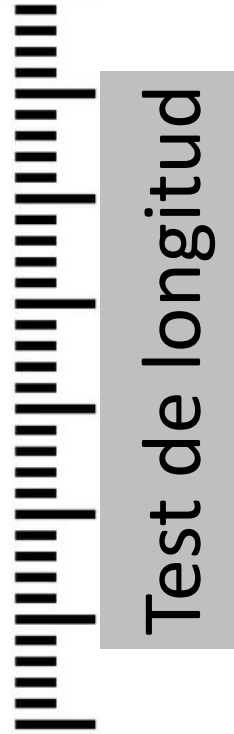
Comparar una cantidad con su respectiva unidad, con el fin de averiguar cuántas veces la segunda está contenida en la primera.

¿Qué longitud tiene este pájaro?



Medir

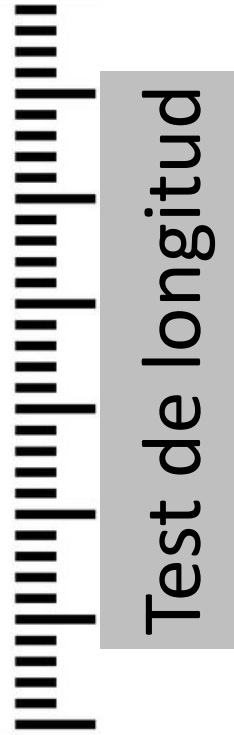
¿Qué longitud tiene este pájaro?



“13 centímetros”

Medir

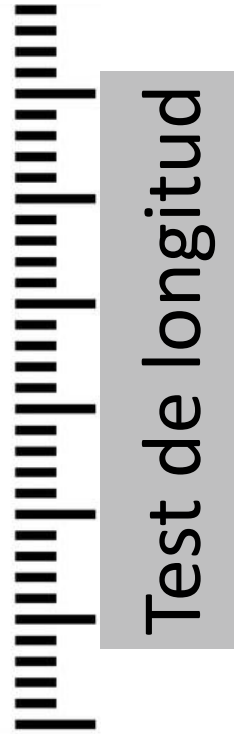
¿Y este otro?



“90 centímetros”

Medir

¿Cuál de los dos tiene mayor longitud?



Águila: 90 cm

“El águila”

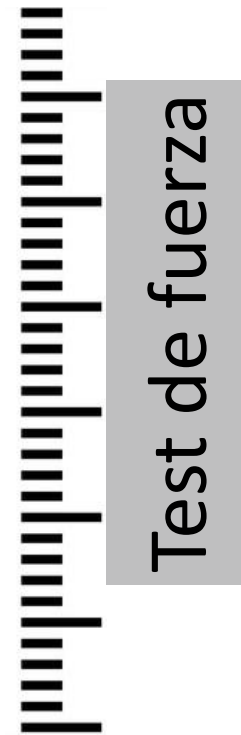
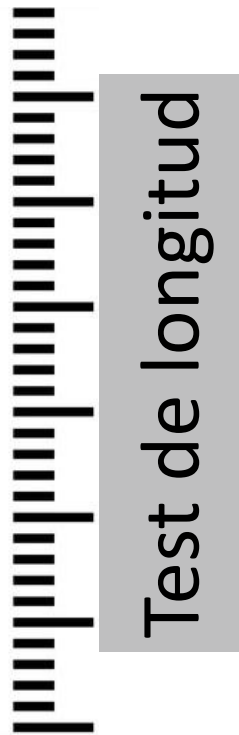
Canario: 13 cm

En Ciencias del Comportamiento y de la Salud medimos....

Inteligencia
Personalidad
Memoria
Atención
Motivación
Afecto
Estrés
Depresión

Percepción
Habilidades sociales
Autoestima
Personalidad patológica
Conocimientos
...

El proceso de medición debe ser...



Fiable: Medida precisa, estable

Válido: Debe poder recabarse evidencia que apoye la manera en que se realiza la medición



CI 70

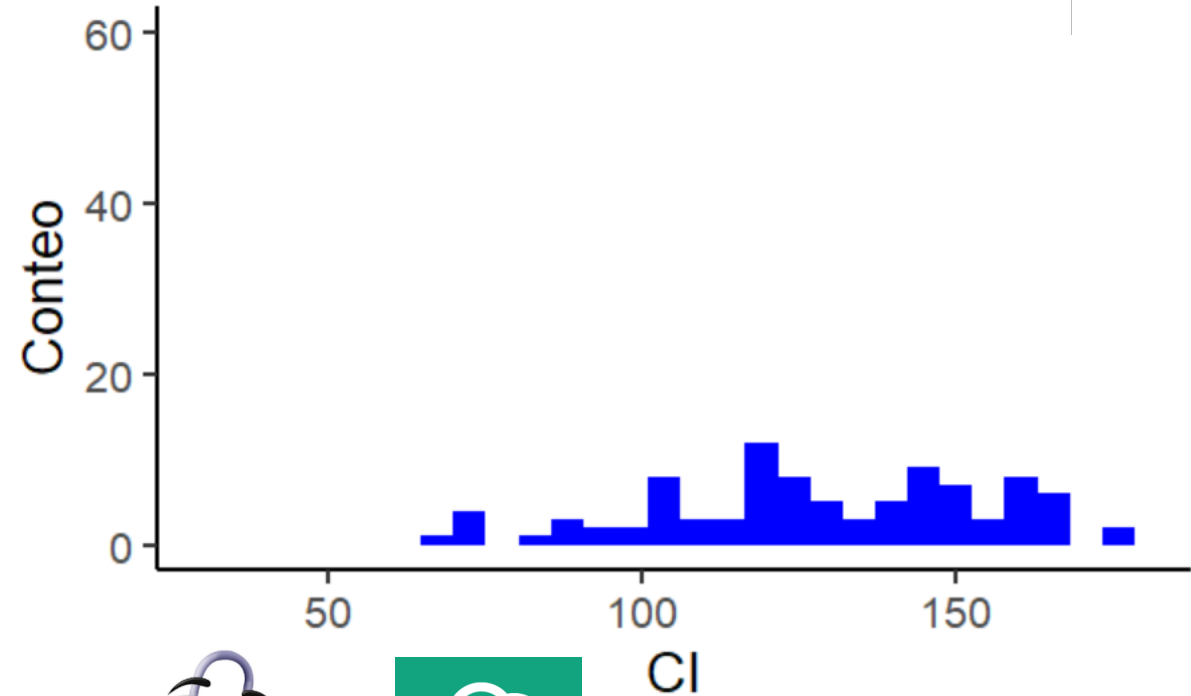


CI 130

Inteligencia
(CI = Cociente Intelectual)

Medimos a ChatGPT 100 veces....

❌ **Fiabile**



CI 70

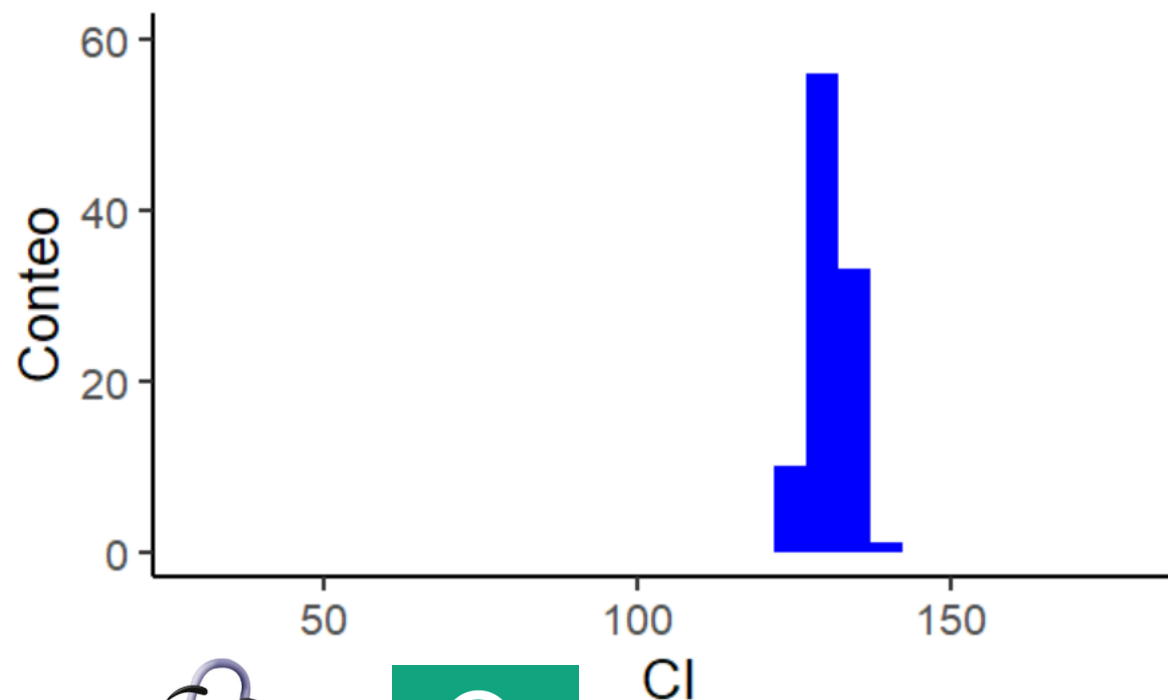


CI 130



Test de inteligencia

Medimos a ChatGPT 100 veces....



CI 70

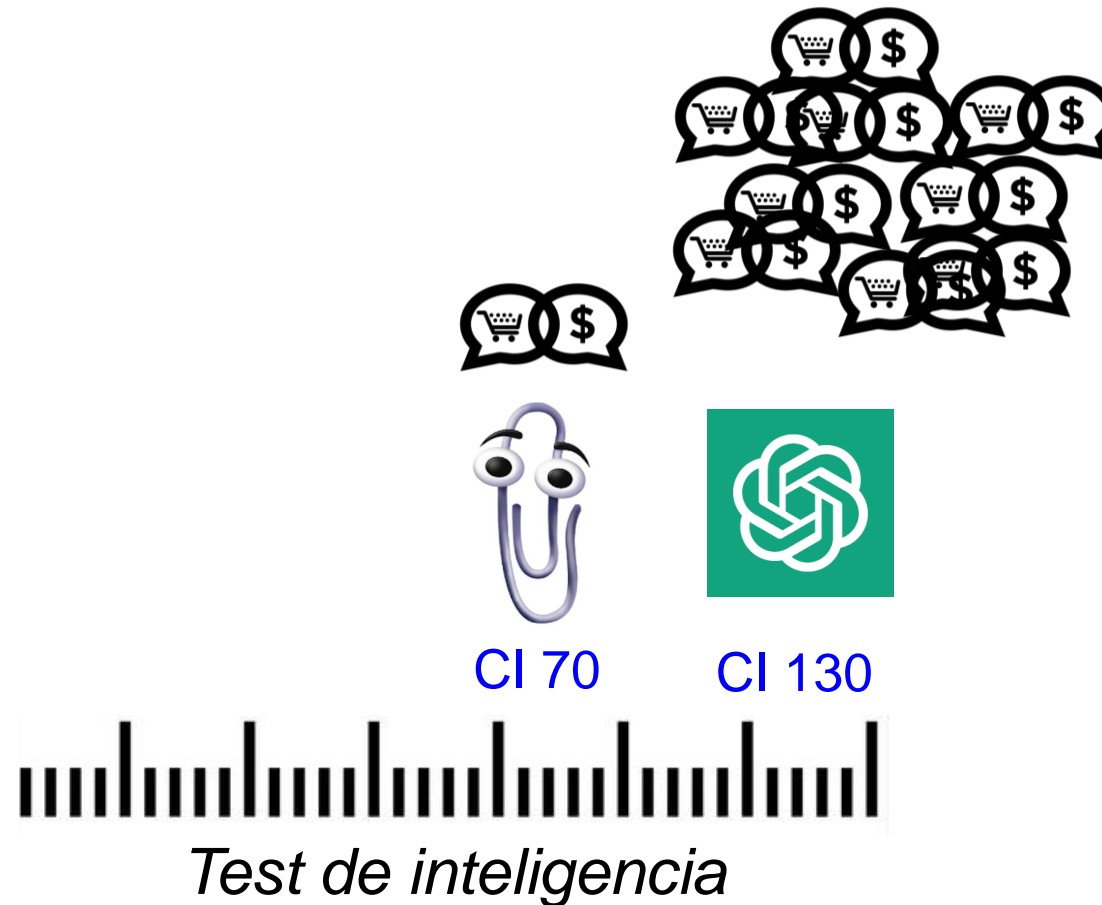


CI 130



Test de inteligencia

Si vemos quien rinde mejor en tareas que requieren inteligencia...



Fiable: Medida precisa, estable

Válida: Debe poder recabarse evidencia
que apoye la manera en que se realiza

Rápida y eficaz



age	irritar	edad	operativo	causante	animo
house	coche	abuelo	doctor	casa	mesa
table	blanco	hogar	mesa	tenis	filete
green	verde	coche	grueso	tema	botella
speed	cantar	velocidad	honor	tinaja	copa
moisture	mitad	seducir	breve	humedad	tramar



⊗	age	irritar	edad	operativo	causante	animo
⊗	house	coche	abuelo	doctor	casa	mesa
⊗	table	blanco	hogar	mesa	tenis	filete
⊗	green	verde	coche	grueso	tema	botella
⊗	speed	cantar	velocidad	honor	tinaja	copa
¿?	moisture	mitad	seducir	breve	humedad	tramar

¿Tiene sentido gastar tiempo preguntando
por *moisture*?

Medición diagnóstica

Medir

¿Cuál de los dos tiene un pico con mayor fuerza?



Test de fuerza

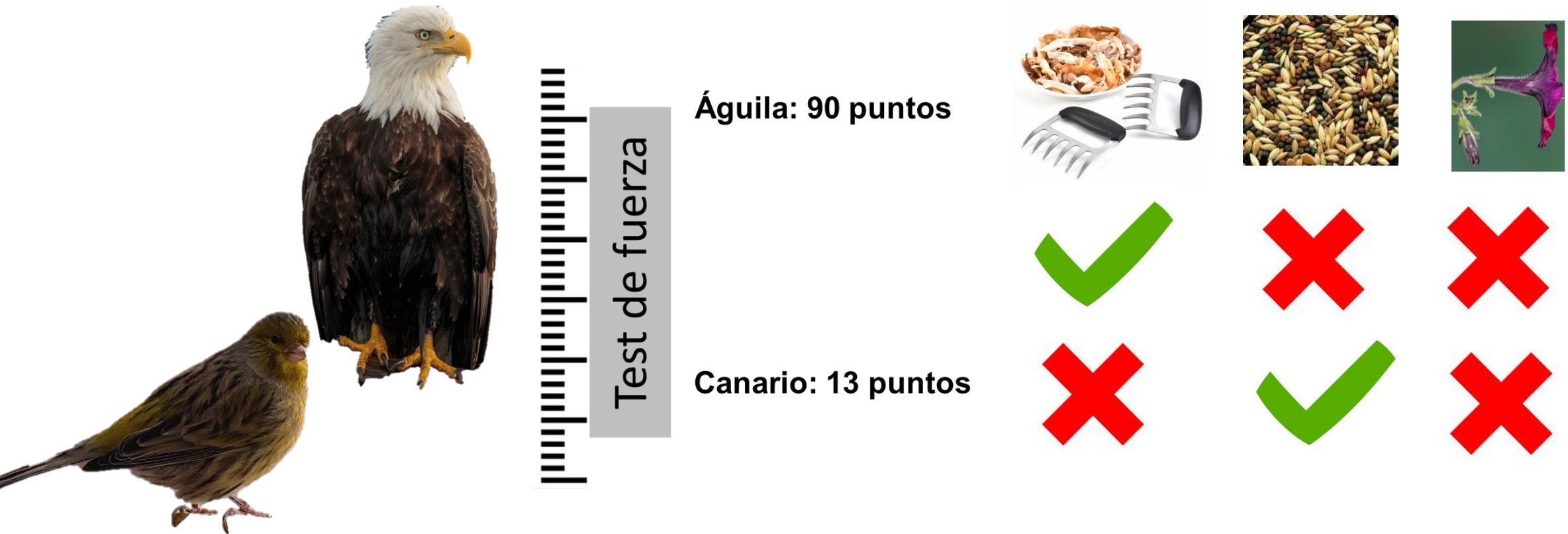
Águila: 90 puntos

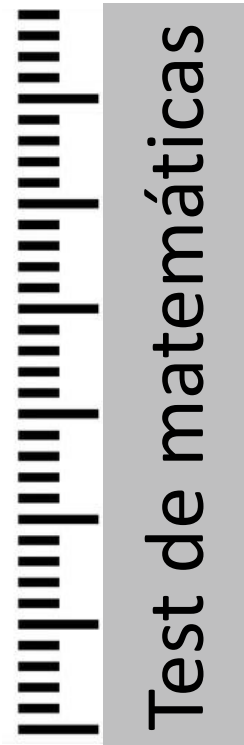
“El águila”

Canario: 13 puntos

Medir

¿Cuál de los dos sirve para comer carne? ¿semillas? ¿néctar?





Persona A: 90 puntos

Persona A: 30 puntos

Persona B: 13 puntos

Sumar y restar
fracciones

Sacar
enteros

Simplificar

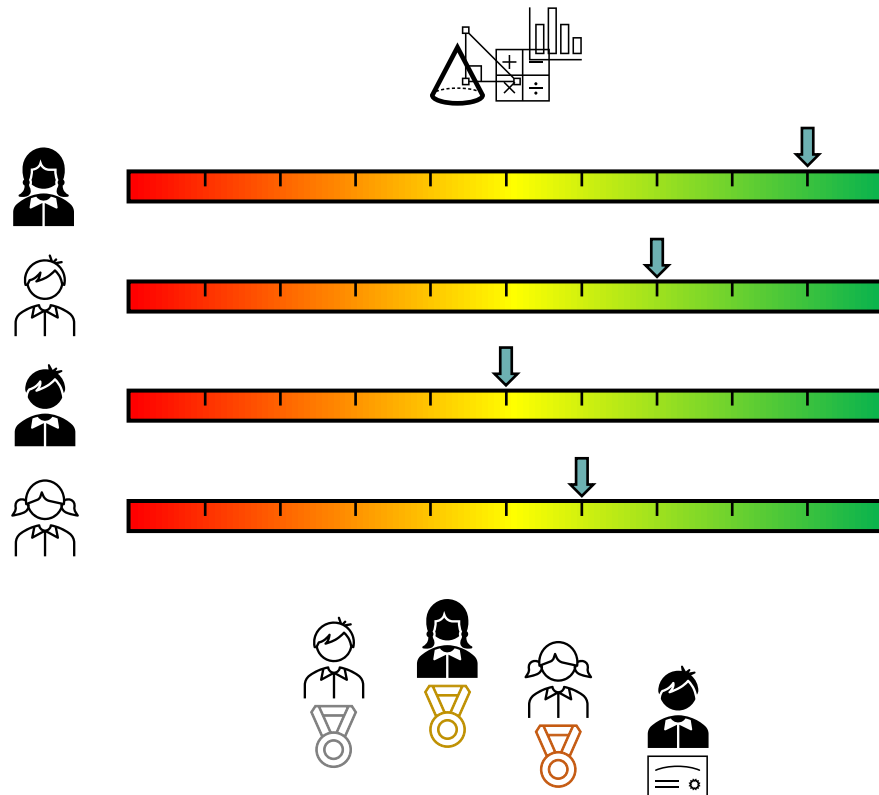


Evaluación sumativa

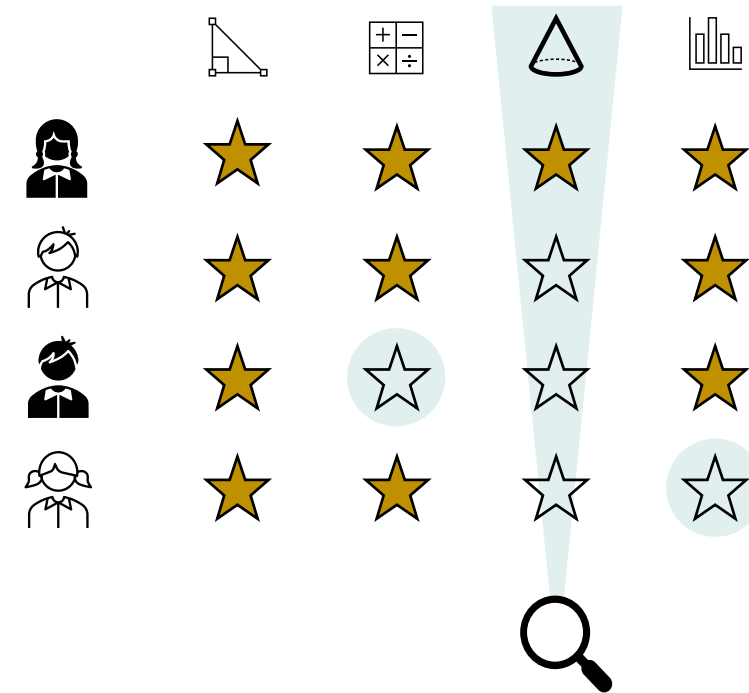
Evaluación formativa

COGNITIVE DIAGNOSIS ASSESSMENTS

Evaluación sumativa



Evaluación formativa



“I am ashamed of the things I’ve done to obtain money for gambling”

DSM-IV-TR - Millon Clinical Multiaxial Inventory-III

He or she has
committed illegal acts

OR

...relied on others to
finance his or her gambling

$$\frac{11}{8} - \frac{1}{8} = ?$$

Competencias Matemáticas

Subtract basic fractions

AND

Reduce to the simplest form

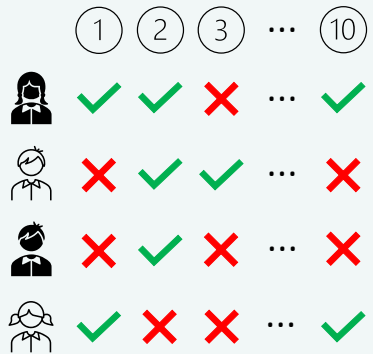
$$\frac{11}{8} - \frac{1}{8} = \frac{10}{8}$$

$$\frac{10}{8} = \frac{5}{4}$$

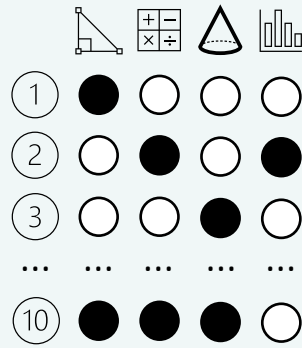
COGNITIVE DIAGNOSIS MODELS

Input

Respuestas



Matriz-Q

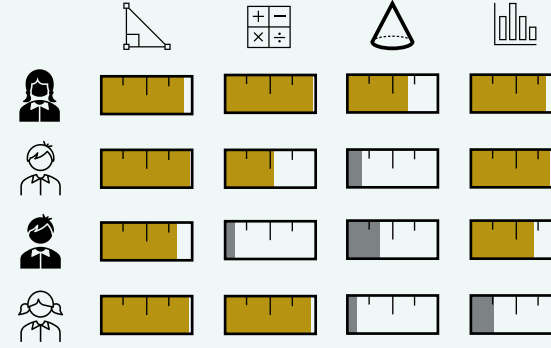


Proceso de respuesta



Output

Probabilidad / Clasificación



Fiabilidad

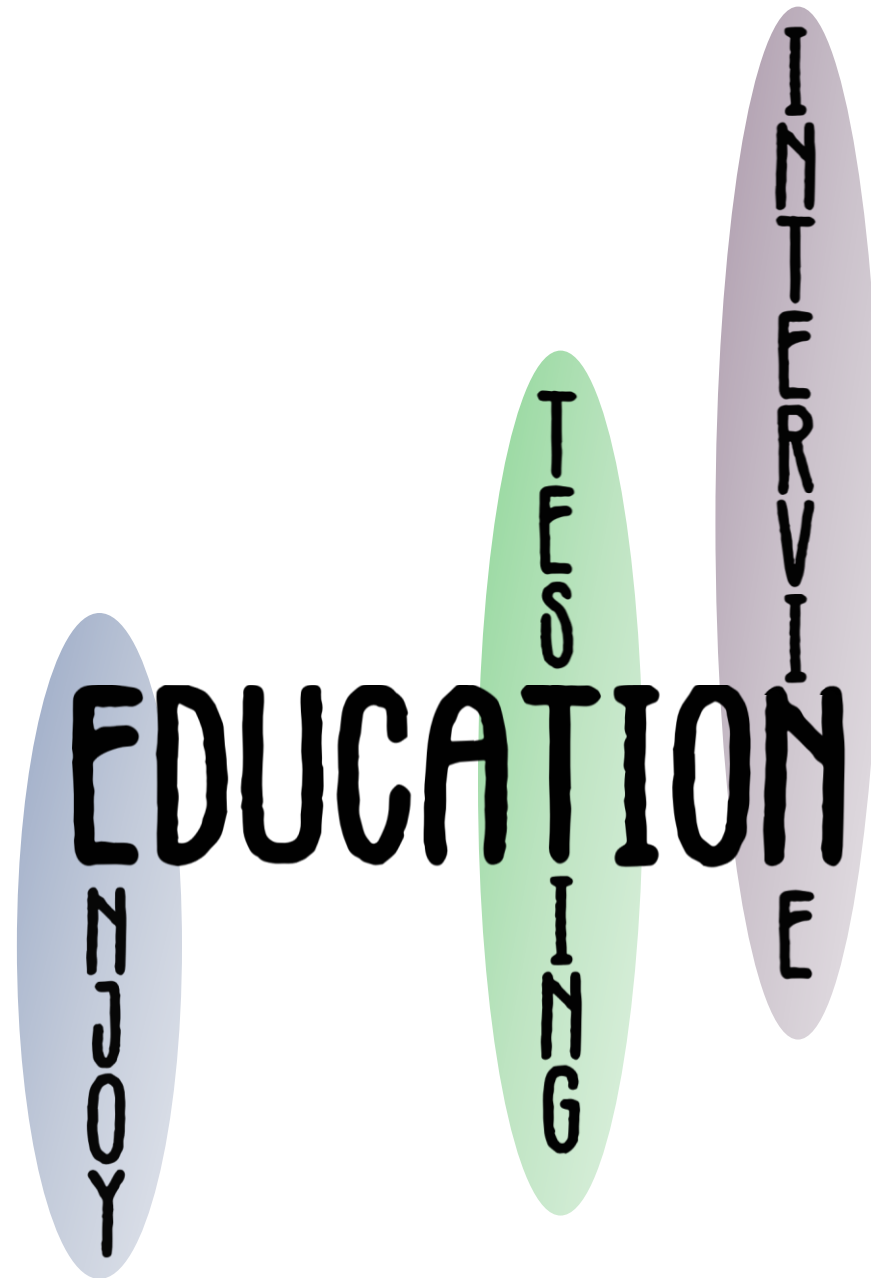
% estimado de
clasificaciones
correctas

Parámetros de los ítems



Ajuste

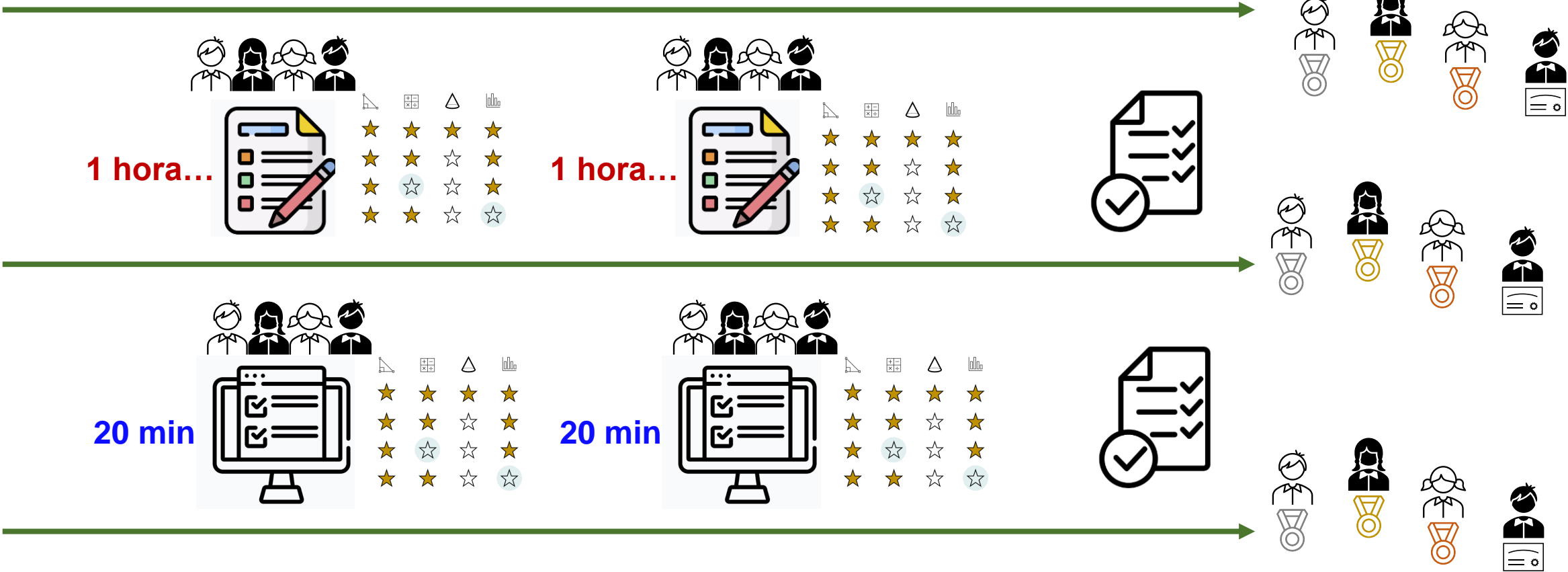
Absoluto y relativo

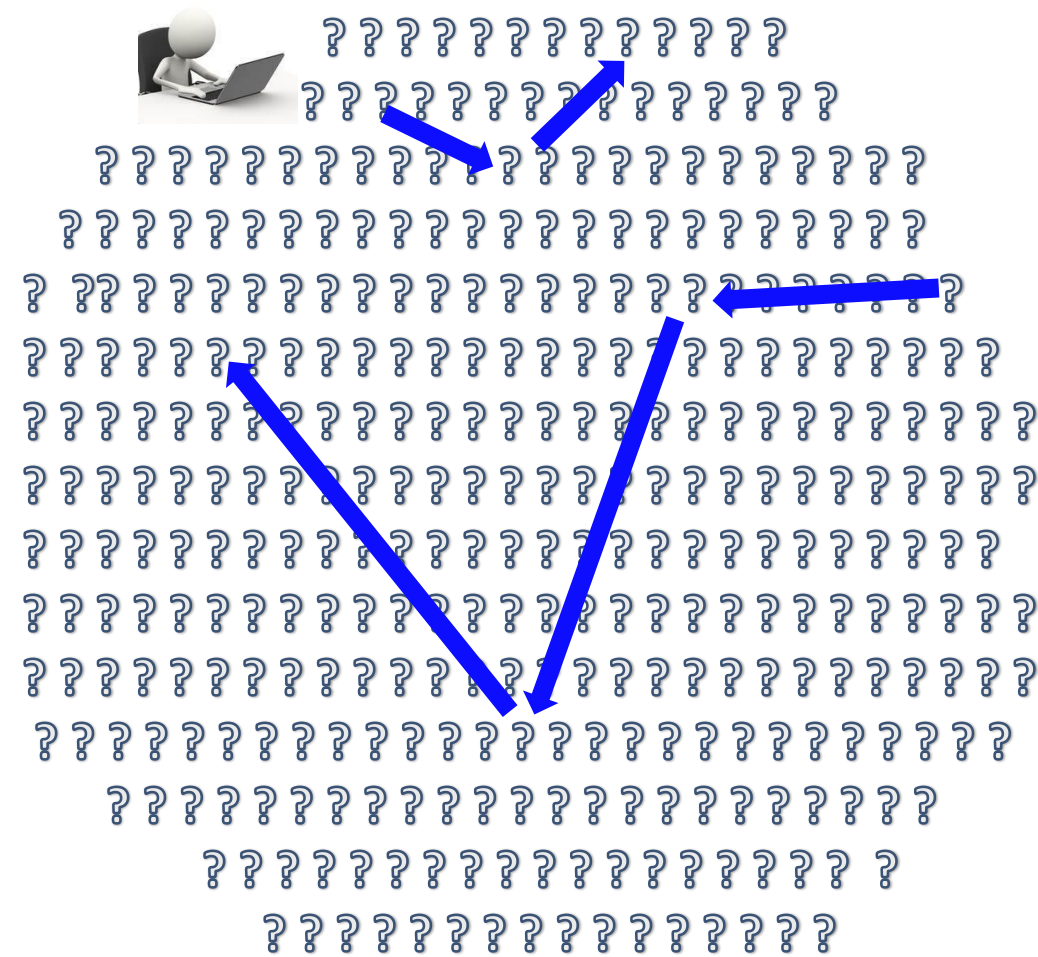


Inicio del curso



Fin





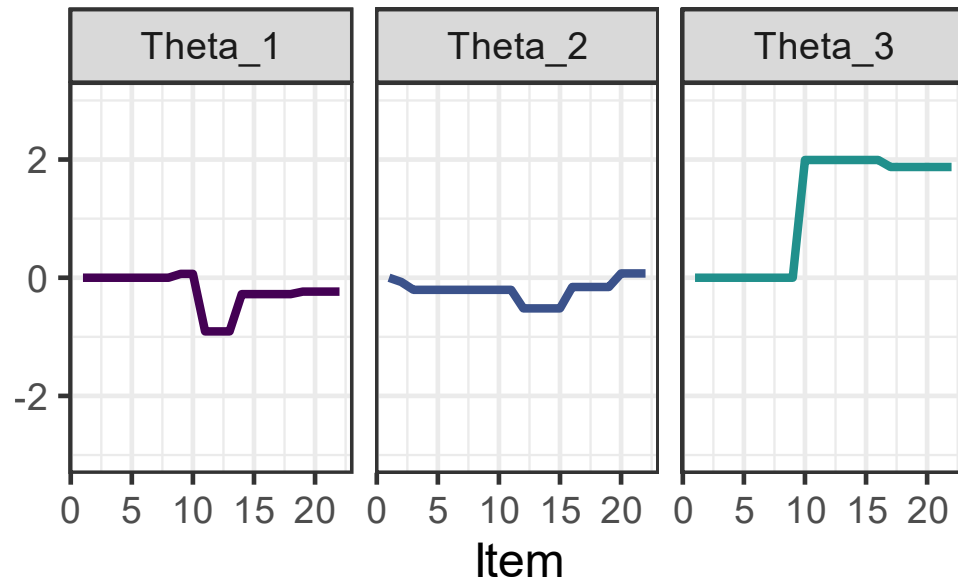
Banco de ítems



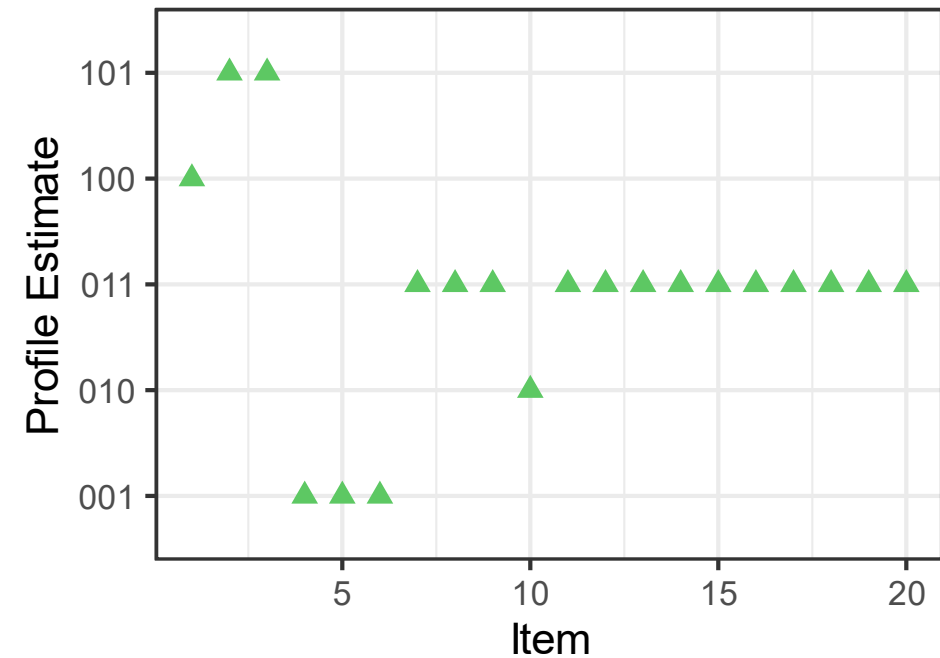
Algoritmo
adaptativo

COGNITIVE DIAGNOSTIC COMPUTERIZED ADAPTIVE TESTING

Test Adaptativo Informatizado Tradicional
(Nivel de rasgo estimado)

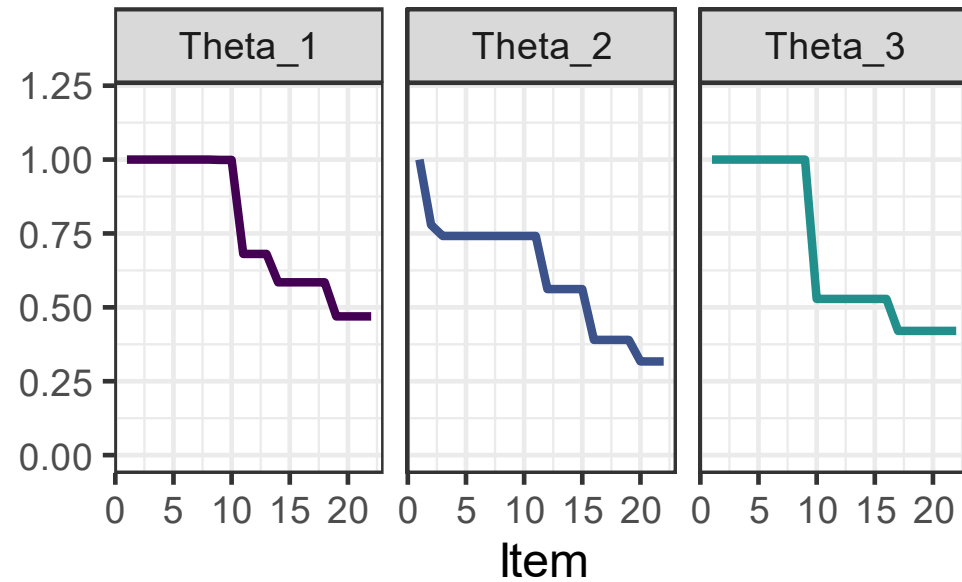


CD-CAT
(Perfil estimado)



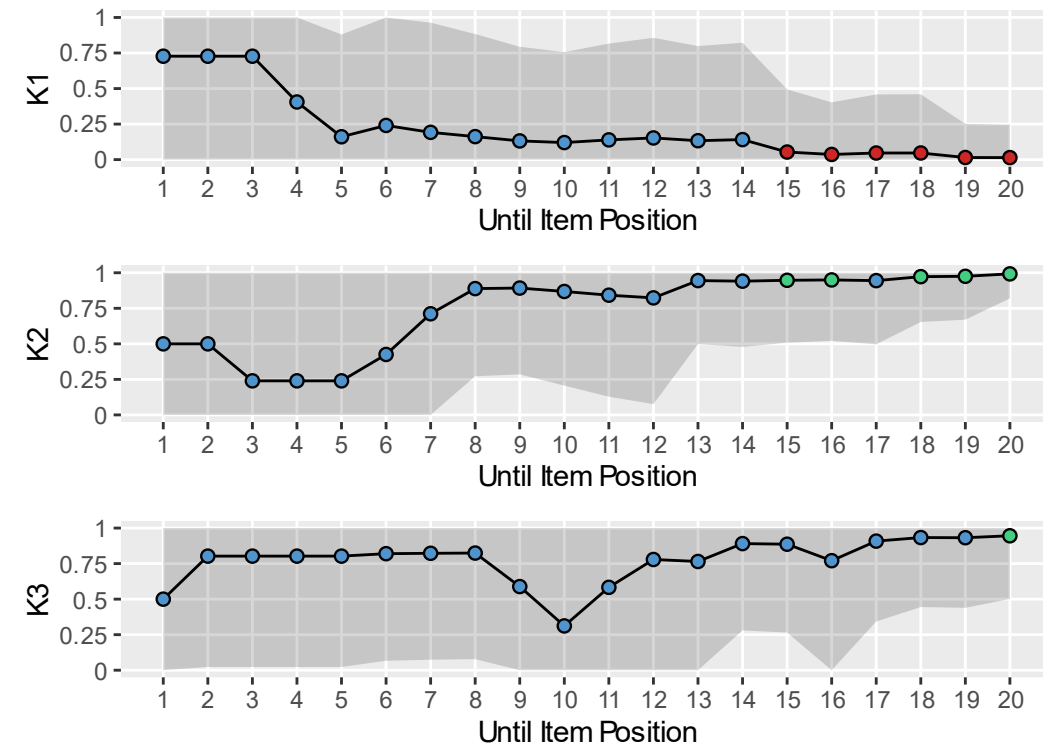
COGNITIVE DIAGNOSTIC COMPUTERIZED ADAPTIVE TESTING

Test Adaptativo Informatizado Tradicional
(Error típico)



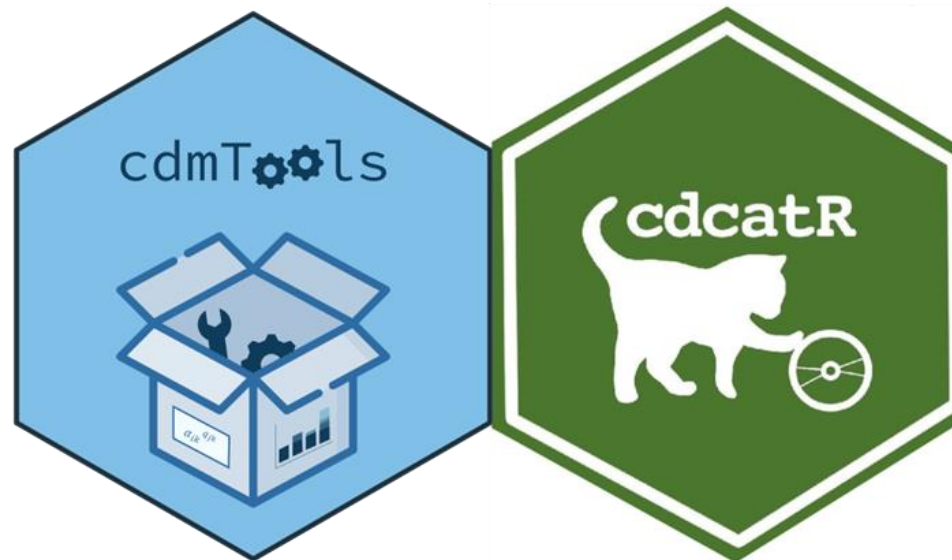
CD-CAT

(Probabilidad posterior)



Paquetes de R

Paquetes “base”: **GDINA** (Ma & de la Torre) y **CDM** (Robitzsch, Kiefer, Cathrice-George & Uenlue)



Medición diagnóstica






Test adaptativos informatizados





Article




Cognitively Diagnostic Analysis Using the G-DINA Model in R

Qingzhou Shi ^{1,*}, Wenchao Ma ¹ , Alexander Robitzsch ^{2,3} , Miguel A. Sorrel ⁴  and Kaiwen Man ¹



Article

cdcatR: An R Package for Cognitive Diagnostic Computerized Adaptive Testing

Miguel A. Sorrel ^{*} , Pablo Nájera  and Francisco J. Abad 

cdmTools

Versión 1.0.4 disponible en CRAN

```
install.packages("cdmTools")  
library(cdmTools)
```

=====

cdmTools Package [Version 1.0.4; 2023-11-10]

More information: <https://github.com/Pablo-Najera/cdmTools>

=====

cdmTools: Funciones

La versión **1.0.4** del paquete incluye las siguientes funciones:

Modelos: **RDINA** (un modelo muy simple) y **GNPC** (más general y complejo)

Evidencias de fiabilidad y validez:

- ¿Mi matriz-Q es adecuada? **is.Qid** (identificación estadística), **paK** y **modelcompK** (dimensionalidad), **estQ** (estimar), **valQ** (validar)
- ¿Las clasificaciones son fiables? **CA.MI**
- ¿Buen ajuste del modelo? **personFit** (ajuste a nivel de sujeto)

Útiles para estudios de simulación: **missQ** (introduce errores en matriz-Q), **orderQ** (ordena matriz-Q)

cdmTools: *RDINA()*

Función principal. Incluye los siguientes argumentos:

INPUT

dat

Q

att.prior

SETTINGS

gate

ESTIMATION (defaults)

est = "Brent"

```
EM.args = list(maxitr = 1000,  
conv.crit = 1e-04, init.phi = 0.2,  
verbose = TRUE),  
tau.alpha = "MAP",  
seed = NULL  
)
```

El output incluye la calibración del modelo RDINA

cdmTools: *RDINA()* (Cont.)

```
#-----Data-----#
```

```
dat <- GDINA::sim30DINA$simdat
```

```
Q <- GDINA::sim30DINA$simQ
```

```
#-----Model estimation-----#
```

```
fit.AND <- RDINA(dat = dat, Q = Q, gate = "AND") # Proceso no compensatorio
```

```
# fit.OR <- RDINA(dat = dat, Q = Q, gate = "OR") # Proceso compensatorio
```

cdmTools: *RDINA()* (Cont.)

fit.AND

=====

R-DINA model

Packages used: cdmTools 1.0.4, NPCD 1.0-11, GDINA 2.9.4

Model estimated using the 'AND' gate and Brent optimization,
000 individuals, 30 items, and 5 attributes.

Estimated $\varphi = 0.2020$

Relative model fit:

-2LL = 35237.27 | npar = 32

AIC = 35301.27 | BIC = 35458.31

CAIC = 35490.31 | SABIC = 35356.68

Estimated classification accuracy (τ):

Test level = 0.698

Profile level = from 0.308 {00000} to 0.958 {11111}

Attribute level = from 0.911 (A2) to 0.916 (A3)

=====

Para gate = "OR":

Relative model fit:

-2LL = 42342.54 | npar = 32

AIC = 42406.54 | BIC = 42563.59

CAIC = 42595.59 | SABIC = 42461.96

cdmTools: RDINA() (Cont.)

```
head(fit.AND$MAP)
```

```
A1 A2 A3 A4 A5 MM
```

```
1 0 1 1 0 1 FALSE # domina los atributos 2, 3 y 5
```

```
2 1 1 0 1 1 FALSE
```

```
3 1 1 0 1 0 TRUE
```

```
4 1 1 0 1 1 FALSE
```

```
5 0 0 1 0 1 FALSE
```

```
6 0 0 0 0 1 FALSE # domina solo el atributo 5
```

```
fit.AND$class.accu$tau_k # clasificaciones fiables (0-1)
```

```
A1 A2 A3 A4 A5
```

```
0.9130451 0.9103373 0.9150374 0.9111614 0.9131903
```

cdcatR

Version 1.0.6 is available in CRAN

```
install.packages("cdcatR")  
library(cdcattR)
```

=====

cdcatR Package [Version 1.0.6; 2022-05-25]

More information: <https://github.com/miguel-sorrel/cdcattR>

=====

cdcatR: Functions

La versión 1.0.6 del paquete incluye las siguientes funciones:

- `cdcat`: Función principal. Lleva a cabo la aplicación CD-CAT
- `cdcat.summary`: Extrae información útil de uno o varios objetos `cdcat`
- `att.plot`: Gráficos de las estimaciones del dominio de atributos
- `gen.itembank`: Genera un banco de ítems
- `gen.data`: Genera datos basados en un objeto `gen.itembank`
- `LR.2step`: Evalúa si G-DINA puede sustituirse por CDM reducidos

También se incluyen algunos conjuntos de datos simulados:

`sim180DINA/sim180GDINA/sim180combination`: Datos, matriz Q y parámetros de los ítems para un banco de 180 ítems con 5 atributos. El modelo generador es DINA/G-DINA/combinación de DINA, DINO y A-CDM

cdcatR: *cdcat()*

Función principal. Incluye los siguientes argumentos:

INPUT

fit
dat
att.prior
initial.distr

SETTINGS

itemSelect
MAXJ
FIXED.LENGTH
precision.cut
startRule
startK
NP.args

CONSTRAINTS

itemExposurecontrol
b
maxr
itemConstraint
constraint.args

MISC.

n.cores
seed
print.progress

La salida contiene el progreso durante el CAT para cada examinado (**est**) y el input (**specifications**)

cdcatR: cdcats() (Cont.)

```
#-----Data-----#
```

```
Q <- simI80GDINA$simQ
```

```
K <- ncol(Q)
```

```
dat <- simI80GDINA$simdat
```

```
att <- simI80GDINA$simalpha
```

```
#-----Model estimation-----#
```

```
fit <- GDINA::GDINA(dat = dat, Q = Q, verbose = 0) # GDINA package
```

```
#fit <- CDM::gdina(data = dat, q.matrix = Q, progress = 0) # CDM package
```

```
#-----CD-CAT-----#
```

```
res.FIXJ <- cdcats(fit = fit, dat = dat, MAXJ = 10)
```

```
res.FIXJ$est[[1]]
```

cdcatR: cdcats() (Cont.)

```
> res.FIXJ$est[[1]]
```

```
$est.cat
```

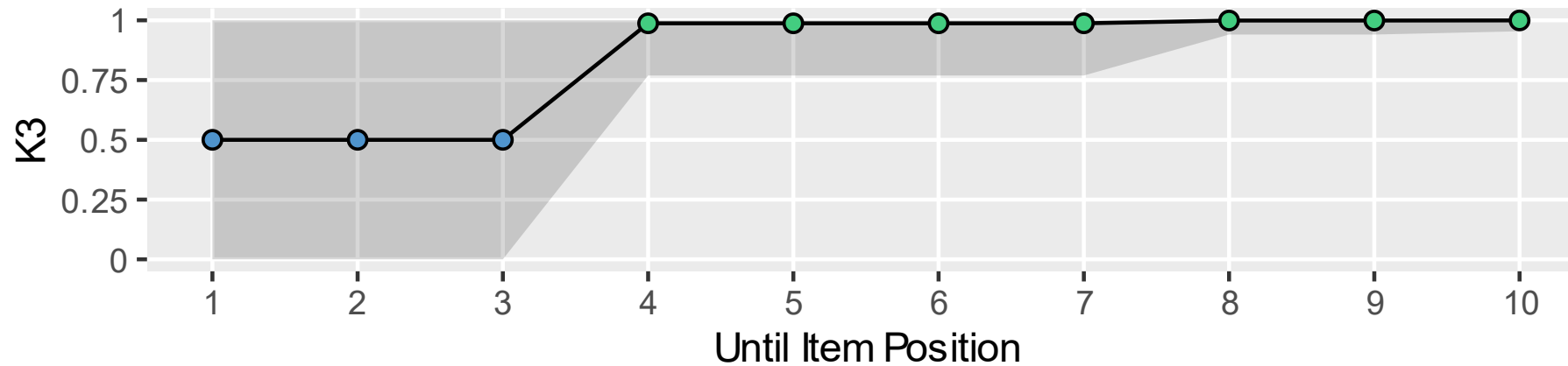
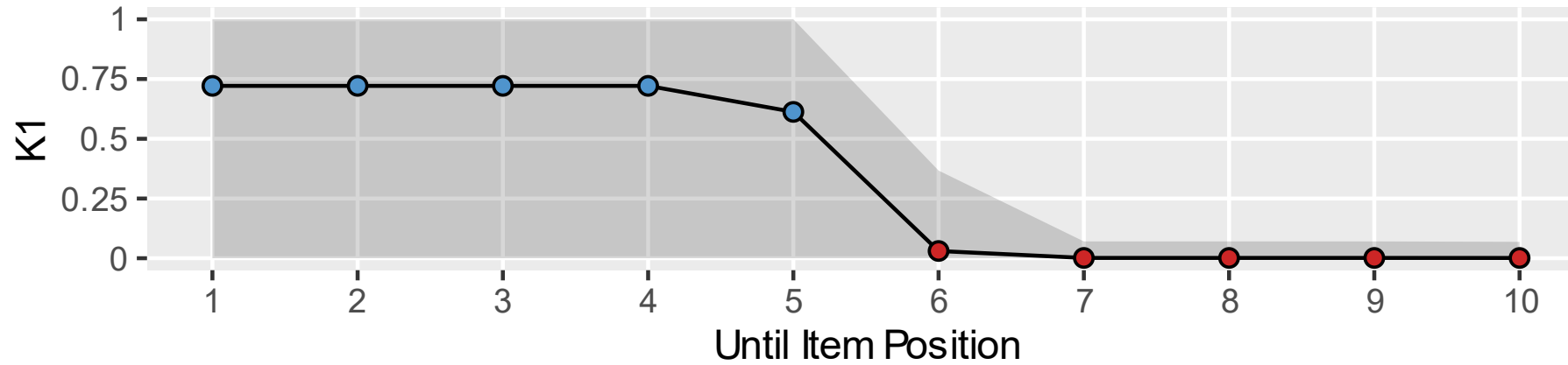
	j	qj	xj	ML	nmodesML	Lik	MAP	nmodesMAP	Post	EAP	K1	K2	K3	K4	K5
1	79	10010	1	10010	8	0.83621	10010	8	0.04688	10010	0.72132	0.5	0.5	0.62383	0.5
2	15	00001	0	10010	4	0.80453	10010	4	0.09375	10010	0.72132	0.5	0.5	0.62383	1e-04
3	12	01000	0	10010	2	0.77668	10010	2	0.1859	10010	0.72132	0.00856	0.5	0.62383	1e-04
4	58	00100	1	10110	1	0.75251	10110	1	0.36712	10110	0.72132	0.00856	0.98742	0.62383	1e-04
5	18	00010	1	10110	1	0.73092	10110	1	0.56784	10110	0.61241	0.00856	0.98742	0.9649	1e-04
6	47	10000	0	00110	1	0.4848	00110	1	0.94228	00110	0.0304	0.00856	0.98742	0.99142	1e-04
7	11	10000	0	00110	1	0.48284	00110	1	0.97061	00110	0.00125	0.00856	0.98742	0.99275	1e-04
8	8	00100	1	00110	1	0.47721	00110	1	0.98209	00110	0.00125	0.00856	0.9991	0.99275	1e-04
9	62	01000	0	00110	1	0.45885	00110	1	0.99046	00110	0.00125	0.00011	0.9991	0.99275	1e-04
10	110	00110	1	00110	1	0.40787	00110	1	0.99749	00110	0.00119	0.00011	0.99946	0.99944	1e-04

```
$item.usage
```

```
[1] 79 15 12 58 18 47 11 8 62 110
```

cdcatR: *att.plot()*

```
att.plot(cdcatt.obj = res.FIXJ, i = 1, k = c(1, 3))
```



cdcatR: *cdcat.summary()*

Información resumida para objetos *cdcat*. Incluye los siguientes argumentos:

cdcat.obj

alpha: Patrones de atributos verdaderos

label: Etiquetas

plots: Mostrar o no gráficos en el output

La salida contiene:

alpha.estimates

item.exposure

alpha.recovery

specifications

cdcatR: `cdcat.summary()` (Cont.)

```
res.FIXJ.sum <- cdc.cat.summary(cdc.cat.obj = res.FIXJ, alpha = att) # % clasificación correcta
res.FIXJ.sum$alpha.estimates$profiles.plot
res.FIXJ.sum$item.exposure$exp.plot
res.FIXJ.sum$alpha.recovery$PCV.plot
```

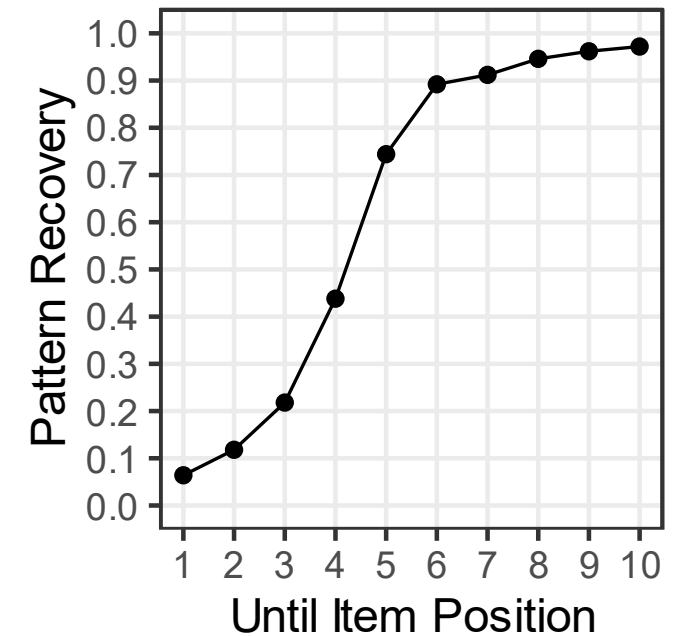
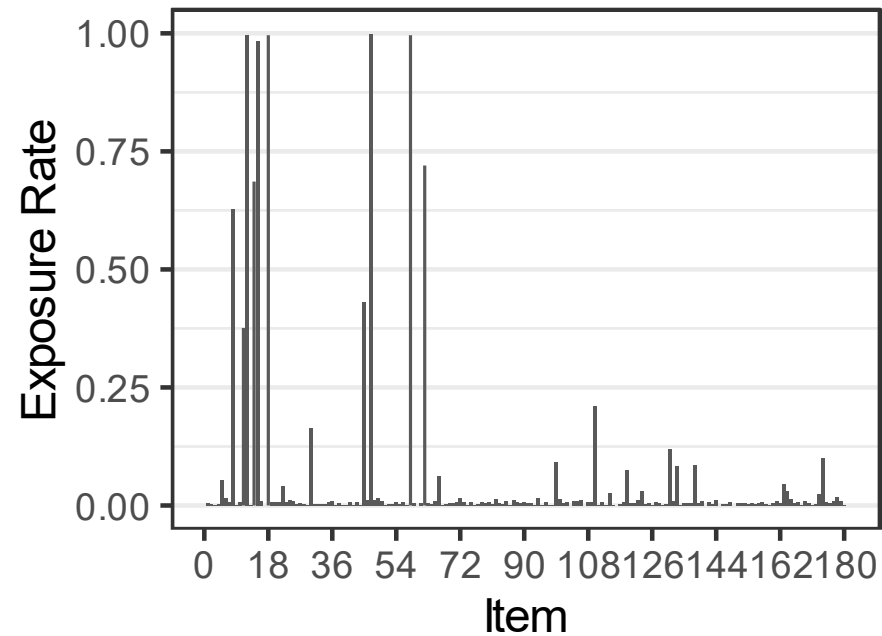
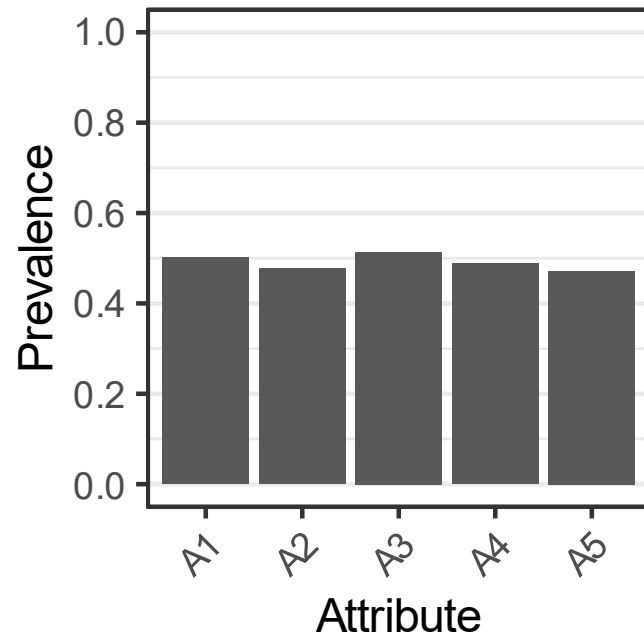


Ilustración: Nonparametric CD-CAT

Yuan-Pei Chang¹, Chia-Yi Chiu¹ and Rung-Ching Tsai²

```
#-----Data Generation-----#
```

```
set.seed(12345)
```

```
Q <- sim180combination$simQ
```

```
bank <- gen.itembank(Q = Q, mean.IQ = .60, range.IQ = .20, model = "DINA")
```

```
simdata <- gen.data(N = 500, R = 2, item.bank = bank)
```

```
dat.c <- simdata$simdat[1:30,,1]
```

```
dat.v <- simdata$simdat[,2]
```

```
att.v <- simdata$simalpha[,2]
```

```
#-----Model estimation-----#
```

```
fitDINA <- GDINA::GDINA(dat = dat.c, Q = Q, model = "DINA") # parametric model
```

Ilustración: Nonparametric CD-CAT (Cont.)

```
#-----CD-CAT-----#
```

```
res.GDI.FIXJ <- cdcats(fit = fitDINA, dat = dat.v,  
  FIXED.LENGTH = TRUE, MAXJ = 25,  
  n.cores = 4, seed = 12345)
```

```
# parametric
```

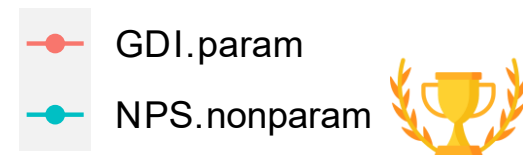
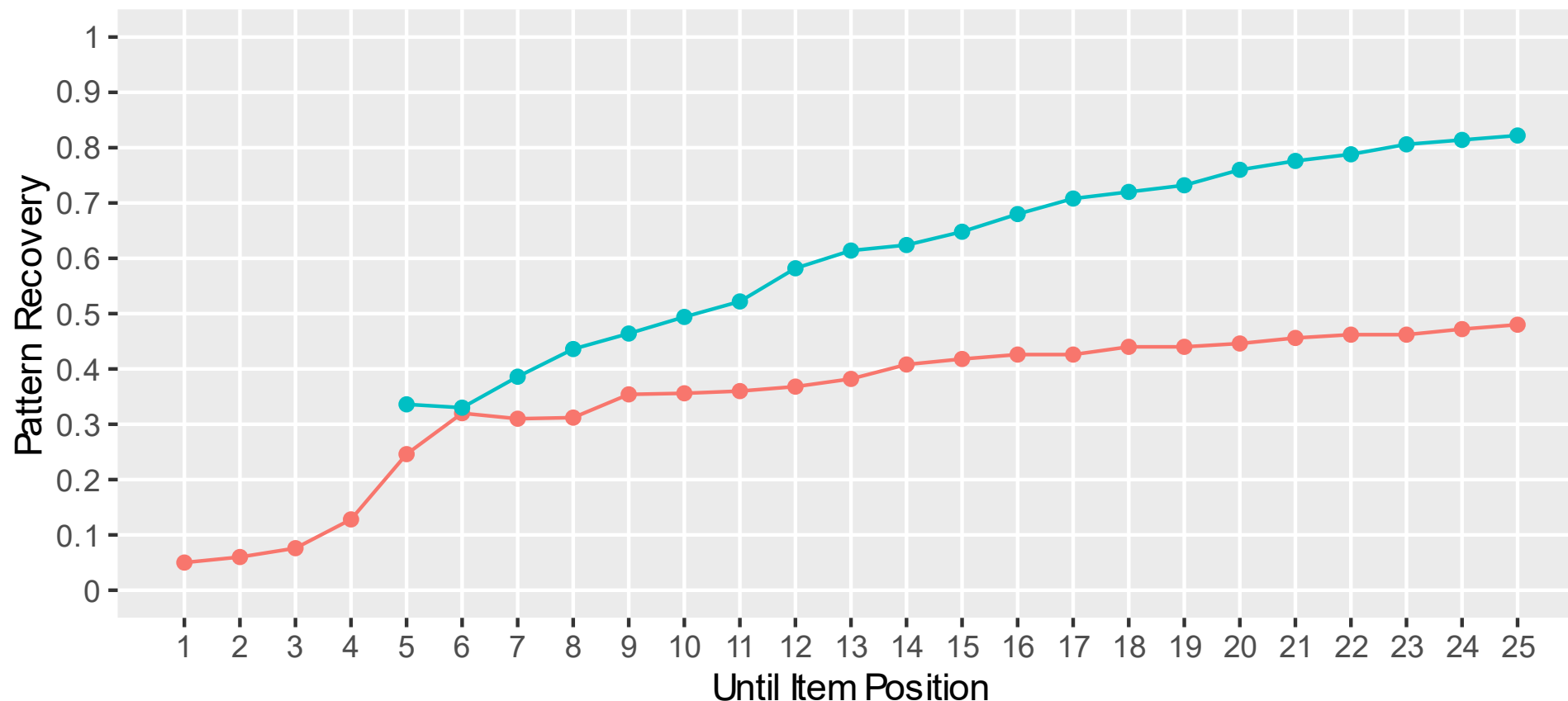
```
res.NPS.FIXJ <- cdcats(dat = dat.v, itemSelect = "NPS",  
  FIXED.LENGTH = TRUE, MAXJ = 25,  
  NP.args = list(Q = Q, gate = "AND", pseudo.prob = TRUE, w.type = 2),  
  n.cores = 4, seed = 12345)
```

```
# nonparametric
```

```
#-----Results-----#
```

```
res.NPS.FIXJ.sum.real <- cdcats.summary(cdcats.obj = list(res.GDI.FIXJ, res.NPS.FIXJ),  
  alpha = att.v, label = c("GDI.param", "NPS.nonparam"))  
res.NPS.FIXJ.sum.real$recovery$PCVcomp
```


Ilustración: Nonparametric CD-CAT (Cont.)



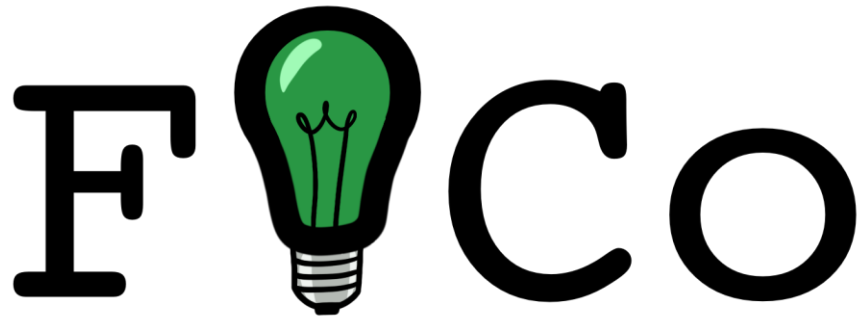
The Restricted DINA Model: A Comprehensive Cognitive Diagnostic Model for Classroom-Level Assessments

Pablo Nájera
Francisco J. Abad
Universidad Autónoma de Madrid

Chia-Yi Chiu
University of Minnesota Twin Cities

Miguel A. Sorrel
Universidad Autónoma de Madrid

Aplicaciones Shiny



<https://psychometricmodelling.shinyapps.io/FoCo/>



<https://psychometricmodelling.shinyapps.io/cdcatS/>

Paso #1: Cargar los datos

[Instrucciones](#)[Datos y Especificaciones](#)[Resultados](#)[Más ▾](#)[Entrada de datos](#)[Especificaciones](#)

Formato del archivo:

Libro de Excel (.xlsx) ▾

- ☒ Incluye nombres de variables.
- ☒ Incluye identificador de sujetos en la primera columna.

Subir archivo:

[Examinar](#)[Plantilla_profe](#)

Upload complete

Base de datos:

Mostrar 10 ▾ registros

	Item.1	Item.2	Item.3	Item.4	Item.5	Item.6	Item.7
1	1	2	3	1	1	3	1
2	1	2	1	2	1	3	3
3	1	2	3	3	3	1	2
4	1	2	9	1	2	3	1
5	1	2	3	2	1	2	2

Paso #2: Especificaciones



Instrucciones

Datos y Especificaciones

Resultados

Más ▾

Universidad Autónoma
de Madrid

Entrada de datos

Especificaciones

Indica el valor para las omisiones:

9

Indica el número de opciones para
cada pregunta:

3

Indica la opción correcta para cada
pregunta:

1,2,3,3,3,1,3,1,1,3

☒ Deseo obtener indicadores adicionales de discriminación
de los distractores.☒ Deseo introducir una matriz Q y realizar la evaluación de
diagnóstico cognitivo.Indica el número de competencias
medidas por la prueba:

3

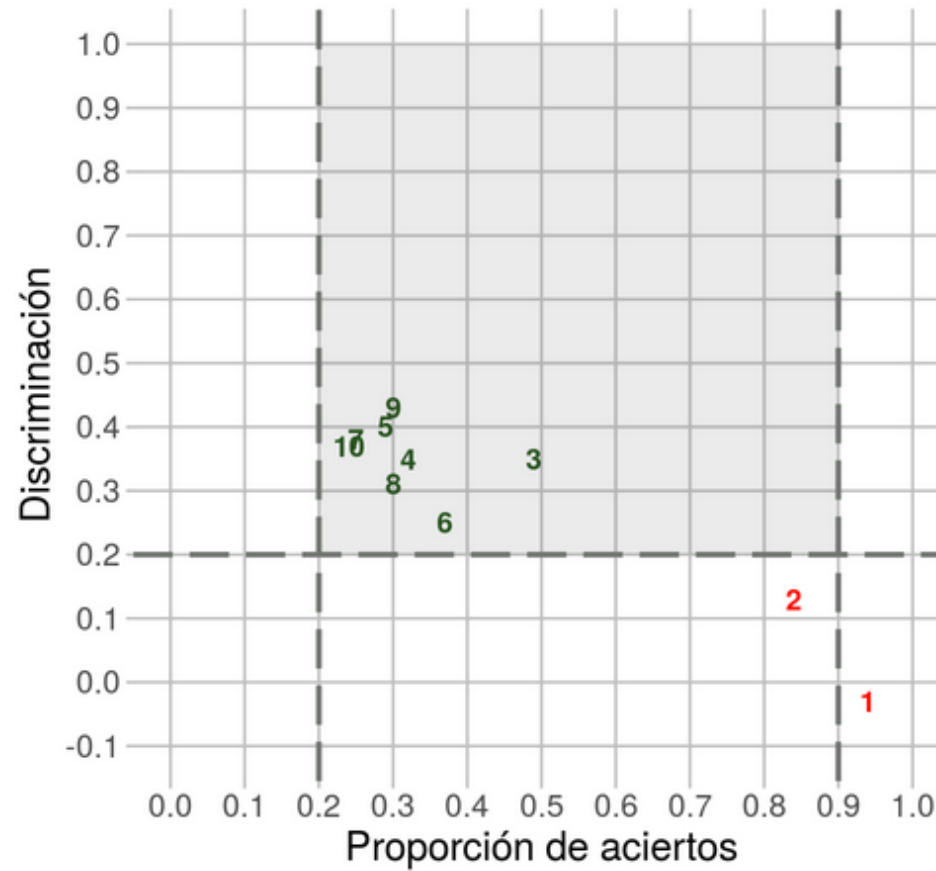
☒ ¿Realizar validación de la matriz Q?
(Solo si hay más de 1 competencia)

Pregunta	C1	C2	C3
Item.1	1	0	0
Item.2	0	1	0
Item.3	0	0	1
Item.4	1	0	1
Item.5	0	1	1
Item.6	1	1	0
Item.7	1	0	1
Item.8	1	1	0
Item.9	0	1	1
Item.10	1	1	1

Analizar

Paso #3: Resultados

Estadísticos para evaluar la calidad de los ítems



Paso #3: Resultados

Gráficos para evaluar el rendimiento de la clase



Paso #3: Resultados

Tabla para evaluar el rendimiento persona a persona

Participante	Aciertos	Errores	Omisiones	Puntuación corregida	C1	C2	C3
001	04	06	00	01.00	0.04	0.64	0.98
002	03	07	00	00.00	0.95	0.02	0.1
003	08	02	00	07.00	0.95	1	1
004	03	06	01	00.00	0.26	0.49	0.04
005	04	06	00	01.00	0.04	0.64	0.98
006	08	02	00	07.00	0.95	1	1
007	04	06	00	01.00	0.17	0.65	0.09
008	05	05	00	02.50	1	0.82	0.01
009	02	07	01	00.00	0.98	0.06	0.02
010	06	04	00	04.00	0.86	0.78	0.54



Crear test

Enter the test ID and password


Enter the test ID and password. To create a new test, click Create CD-CAT. If you want to check an existing, already registered CD-CAT, click on Check CD-CAT.


Test ID

PSICOMETRIA

Test Password

•••••

 Create CD-CAT

 Check CD-CAT

Test created successfully.

itembank.xls

Item.No.	Item	Correct	Option.1	Option.2	Option.3
1	How many oceans are there on Earth?	5	4	5	6
2	What animal is the fastest?	Cheetah	Cheetah	Lion	Gazelle
3	Which of the following is one of the Big Five animals?	Lion	Crocodile	Whale	Lion
...

Qmatrix.xls

Item	Earth	Space	Animals
1	1	0	0
2	0	0	1
3	0	0	1
...



Crear test

Select specifications and upload item bank

Here you can select the specifications of the CD-CAT.

Number of items to administer

Item response process

Conjunctive

Here you can upload an Excel file with the item bank and the Q-matrix specification (check Instructions to see the required format).

Upload item bank

Browse...

itembank.xlsx

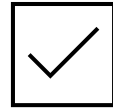
Upload complete

Upload Q-matrix

Browse...

Qmatrix.xlsx

Upload complete



Hacer test

Enter the test and your ID


Enter the test ID and your ID. If both ID have been already entered before, the app will show your test results. Otherwise, the test will be administered to you.

Test ID

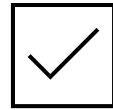
PSICOMETRIA

Your ID

Lord

 Submit

This is the first time you access to this test. You can take the test in the following.



Hacer test

Your responses to the items				Mastery of the measured attributes	
Item	Correct	Attributes	Answer	Attribute	
En el ámbito de la Psicometría, ¿cuál de los sigui	La Teoría de Respuesta al Ítem.	Teoría	Correct		
Un test está formado por 4 ítems dicotómicos. No s	Item 2.	Cálculo	Correct		
Si en un modelo unifactorial un ítem recibe un pes	0,51	Cálculo	Correct		
El coeficiente alfa de un test de 10 ítems es 0,85	Los ítems, en general, covarían alto y positivamente entre sí.	Teoría	Correct		
En un test, sabiendo que la varianza de las puntua	0,66 (=1- 20/60)	Cálculo	Correct		
En el caso de tener 2 factores independientes entr	iguales que los pesos de estructura.	Teoría	Correct		
Los coeficientes de fiabilidad y de validez de un	No es posible alcanzar un coeficiente de validez de 0,60 alargando el test	Cálculo	Correct		
En el análisis exhaustivo de un ítem se calcula la	-0,33	Teoría	Correct		
Tenemos 3 test: (1) un test A, (2) el test que res	10	Cálculo	Correct		
En test de rendimiento típico, la recodificación e	hace que tenga sentido sumar los ítems en una sola puntuación.	Teoría	Correct		



Ver resultados

Check specifications and download results

Here you can check the specifications of the CD-CAT.

Number of items to administer: 10
Item response process: Conjunctive
Number of current responses: 9

Here you can check and download the current results, as well as the item bank and Q-matrix specification.

[👁 Check current results](#)[📄 Download current results](#)[📄 Retrieve item bank](#)[📄 Retrieve Q-matrix](#)

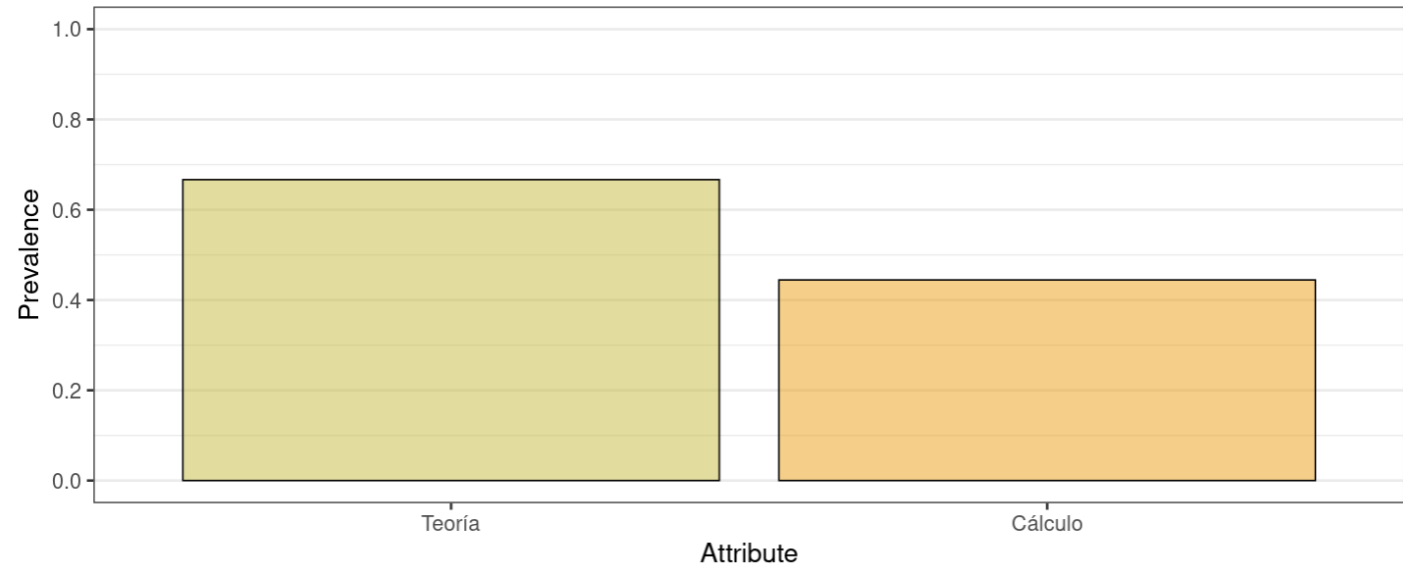


Ver resultados

Individual results

ID	Teoría	Cálculo	Prob.Teoría	Prob.Cálculo
Lord	1	1	0.99661	0.99661
Samejima	1	1	0.99661	0.99661
Spearman	1	1	0.99661	0.99661
Tribbiani	0	0	0.00932	0.00932
Slmpson	0	1	0.02601	0.97399
Bandura	1	0	0.97399	0.02601
Azula	0	0	0.02601	0.02601
Connor	1	0	0.99068	0.00932
Smith	1	0	0.87764	0.12236

Aggregated results



Ilustración

Aplicación a una Asignatura

Definición de **Q** y generación de ítems



Grado en Psicología
Psicometría

Tema
T1: Introducción a la Psicometría
T2: Construcción de un test y análisis de ítems
T3: Modelo clásico y fiabilidad
T4: Concepto y evidencias de validez
T5: Análisis factorial
T6: Interpretación de puntuaciones
T7: Introducción a la TRI

Tema	Componentes	Contenido	Tiempo dedicado (1-3)	Teoría (0-1)	SPSS (0-1)
T1: Introducción a la Psicometría	T1.A	Antecedentes históricos	1	1	0
	T1.B	Tipos de test	2	1	0
T2: Construcción de un test y análisis de ítems	T2.A	Diseño del test	2	1	0
	T2.B	Redacción: Óptimo y típico	3	1	0
	T2.C	Análisis: Cuantificación	3	1	1
	T2.D	Análisis: Media y varianza	3	1	1
	T2.E	Análisis: Discriminación y validez	3	1	1
	T2.F	Análisis: Distractores	3	1	1
	T2.G	Análisis: Puntuación corregida	3	1	1
T3: Modelo clásico y fiabilidad	T3.A	Supuestos TCT	2	1	0
	T3.B	Fiabilidad: Concepto	3	1	0
	T3.C	Fiabilidad: Test alargado	3	1	0
	T3.D	Fiabilidad: Aproximaciones	3	1	1
	T3.E	Fiabilidad: Error típico de medida	2	1	0
	T3.F	Fiabilidad: Factores que influyen	2	1	0
T4: Concepto y evidencias de validez	T4.A	Validez: Concepto	2	1	0
	T4.B	Validez: Evidencias contenido	3	1	0
	T4.C	Validez: Convergente/discriminante	3	1	0
	T4.D	Validez: Criterio			
T5: Análisis factorial	T5.A	Conceptos básicos	2	1	0
	T5.B	Número de factores	3	1	1
	T5.C	Rotación	2	1	1
	T5.D	Puntuaciones factoriales	1	1	0
T6: Interpretación de puntuaciones	T6.A	Centiles	3	1	1
	T6.B	Puntuaciones típicas y escalas derivadas	2	1	0
	T6.C	Baremos cronológicos	1	1	0
	T6.D	Puntos de corte: Angoff y Nedelsky	1	1	0
T7: Introducción a la TRI	T7.A	TCT vs. TRI	1	1	0
	T7.B	Modelos logísticos	1	1	0
	T7.C	Supuestos	1	1	0

COMPETENCIAS

C1:
Teoría

C2:
Cálculo

C3:
Tablas y Figuras

PROCESOS

P1:
Recordar

P2:
Comprender

P3:
Aplicar

Tema		C1: Teoría	C2: Cálculo	C3: Tablas y figuras	P1: Recordar	P2: Comprender	P3: Aplicar
T3	T3.A.1 (1)	1	0	1	0	1	0
	T3.B.1 (1)	0	1	0	1	0	0
	T3.C.1 (1)	0	1	0	0	1	0
	T3.C.2 (1)	0	1	0	0	0	1
	T3.C.3 (1)	1	1	1	0	0	1
	T3.D.1 (1)	1	0	0	1	0	0
	T3.D.2 (1)	1	0	1	0	1	0
	T3.E.1-2 (2)	1	0	0	0	1*	0*
	T3.F.1-2 (2)	1	0	0	1	0	0
Total ítems 11		8	4	3	4	5*	2*

T3.A.1 (1)		
C1:	Teoría	1
C2:	Cálculo	0
C3:	Tablas y figuras	1
P1:	Recordar	0
P2:	Comprender	1
P3:	Aplicar	0

1. A1. A continuación, se muestran las puntuaciones de 5 estudiantes en dos formas paralelas se muestran a continuación. En la tabla los valores de *V* y *E* se muestran en cursiva para indicar que estos datos no se conocen nunca en la realidad:

	X1	<i>V1</i>	<i>E1</i>	X2	<i>V2</i>	<i>E2</i>
Estudiante 1	10	<i>10</i>	<i>0</i>	10	<i>10</i>	<i>0</i>
Estudiante 2	12	<i>12</i>	<i>0</i>	12	<i>12</i>	<i>0</i>
Estudiante 3	18	<i>16</i>	2	18	<i>16</i>	2
Estudiante 4	14	<i>16</i>	-2	16	<i>16</i>	<i>0</i>
Estudiante 5	16	<i>16</i>	<i>0</i>	14	<i>16</i>	-2

La tabla de correlaciones entre las variables es:

	X1	V1	E1	X2	V2	E2
X1	1,00					
V1	0,89	1,00				
E1	0,45	0,00	1,00			
X2	0,90	0,89	0,22	1,00		
V2	0,89	1,00	0,00	0,89	1,00	
E2	0,22	0,00	0,50	0,45	0,00	1,00

En relación con los supuestos de la TCT, en estos datos no se cumple:

- El primer supuesto (la puntuación observada la suma de la puntuación verdadera más un error de medida).
- El segundo supuesto (el valor esperado del error de medida es igual a 0).
- El cuarto supuesto (los valores de error de medida de dos tests distintos no correlacionan)

Aplicación a una Asignatura

Banco de Ítems

- Un total de 85 ítems diseñados
- Cinco personas expertas en Psicometría
- Cada atributo se mide un mínimo de 20 veces
- 100% de las combinaciones de atributos representadas en **Q**

Recogida de datos

- 5 grupos de Psicometría UAM
- 7 autoevaluaciones disponibles en Moodle al finalizar cada tema
- 366 respuestas recibidas de un total de 237 estudiantes (89% matrícula)
- A efectos de esta presentación se analizar la peor respuesta en caso de múltiples intentos

Análisis

- **GDINA**: Estimación, ajuste, fiabilidad
- **cdmTools**: Dimensionalidad, validación de **Q**, fiabilidad (imputación múltiple)

Aplicación a una Asignatura

Resultados

- Ajuste aceptable:
 - Q-competencias AIC 19037.49, RSMEA 0.0585 90%-CI: [0.0561, 0.0608],
4 pares de ítems problemáticos
 - Q-procesos AIC 18978.83, RSMEA 0.0613 90%-CI: [0.0590, 0.0636],
1 par de ítems problemático
- Fiabilidad adecuada:
 - Q-competencias tau = 0.9206 (imputación múltiple: 0.7179)
 - Q-procesos tau = 0.9044 (imputación múltiple: 0.8950)

Autoevaluaciones: Temas 1-3 (Grupo 3)

Grupo	G3	Total
N.º estudiantes	36	209
Total Media (SD)	20.39 (8.12)	24.13 (7.40)

Estimaciones CDM

Competencias

T (Preguntas teóricas)
C (Preguntas con cálculos)
F (Preguntas con tablas/figuras)

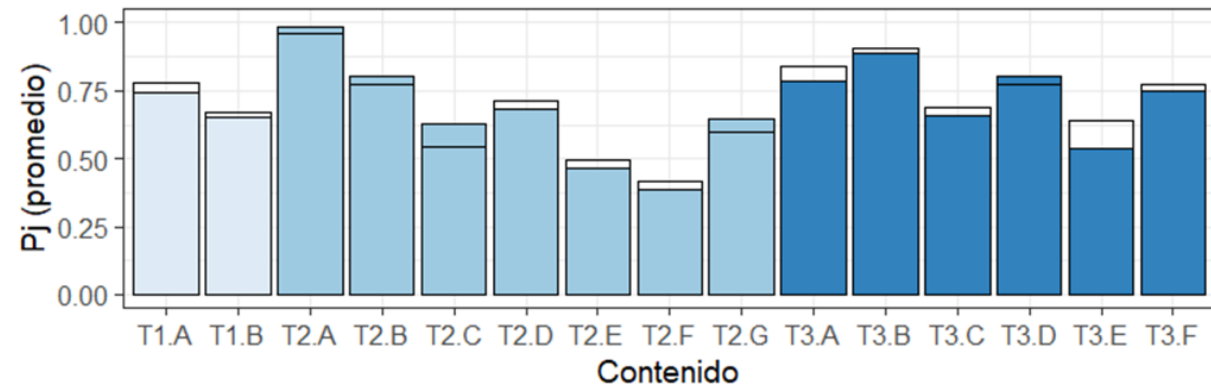
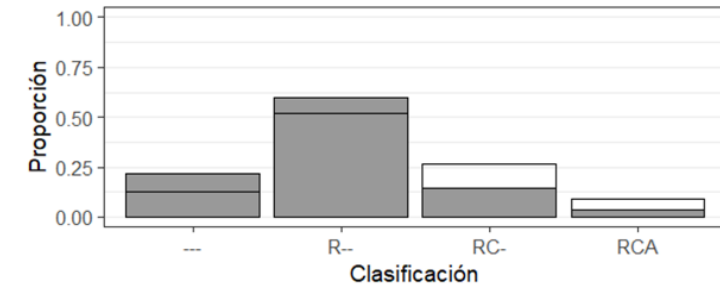
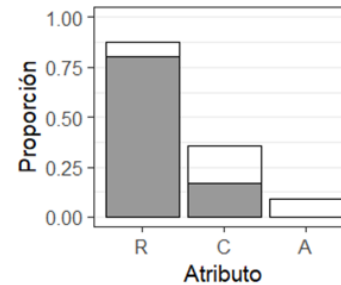
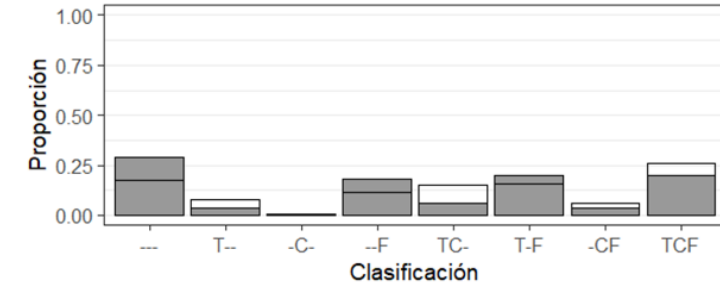
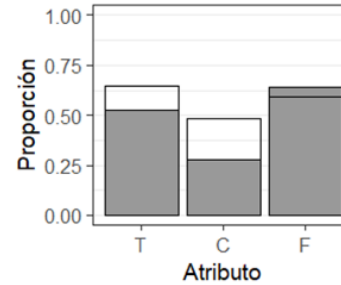
Procesos

R (Recordar)
C (Comprender)
A (Aplicar)

Correlación promedio de atributos = 0.31

Fiabilidad adecuada en todas las clasificaciones

T1.A	Antecedentes históricos
T1.B	Tipos de test
T2.A	Diseño del test (QUÉ, A QUIÉN, PARA QUÉ)
T2.B	Redacción: Óptimo y típico
T2.C	Análisis: Cuantificación
T2.D	Análisis: Media y varianza
T2.E	Análisis: Discriminación y validez
T2.F	Análisis: Distractores
T2.G	Análisis: Puntuación corregida
T3.A	Supuestos TCT
T3.B	Fiabilidad: Concepto
T3.C	Fiabilidad: Test alargado
T3.D	Fiabilidad: Aproximaciones
T3.E	Fiabilidad: Error típico de medida
T3.F	Fiabilidad: Factores que influyen



Estudiante:

Este informe tiene por objetivo facilitarte información que te permita conocer mejor tu dominio de los diferentes aspectos de la asignatura de Psicometría. La información que se incluye aquí está basada en tus respuestas a las tareas de autoevaluación y no se usará para puntuar en la asignatura. Sin embargo, conocer los fortalezas y debilidades puede ayudarte a organizar el estudio y, por consecuencia, tener un mejor rendimiento en la asignatura.

Competencias Evaluadas

En las tareas de autoevaluación a las que has respondido, hemos buscado conocer tu aprendizaje en la asignatura en dos aspectos diferentes: 1) con respecto a los contenidos de la asignatura en sí, y 2) con respecto a qué eres capaz de hacer con el contenido que has aprendido. En otras palabras, hemos buscado conocer no solo si puedes recordar de forma general el contenido del temario, pero también si lo has podido entender bien y si sabrías aplicarlo para resolver problemas en el futuro.

Así, hemos agrupado los contenidos del temario en 3 categorías, que son:

Contenidos del temario

Teoría: Implica que sepas el contenido teórico de la asignatura. Por ejemplo, que conozcas aspectos históricos de la Psicometría, así como las diferentes teorías expuestas en clase.

Cálculo: Significa que domines las diferentes fórmulas y relaciones numéricas de la asignatura. Por ejemplo. . .

Tablas y Figuras: Implica que sepas utilizar las tablas y figuras relevantes en Psicometría. Por ejemplo, que sepas interpretar correctamente las tablas de frecuencias de respuesta de diferentes sujetos a un test.

Por otro lado, los 3 niveles de habilidades con el contenido son:

Habilidades con el contenido

Recordar: Implica que seas capaz de reconocer, definir, enumerar, o describir los contenidos de la asignatura.

Comprender: Representa que, además de recordar, puedas interpretar, ejemplificar, clasificar, resumir, inferir, o explicar los contenidos.

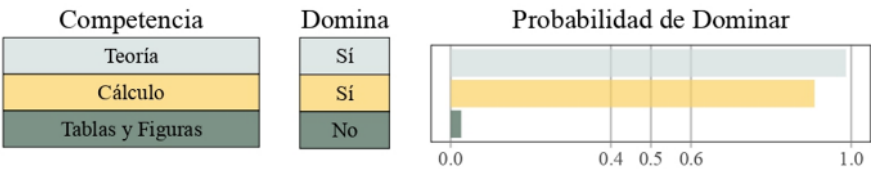
Aplicar: Hace referencia a que, además de recordar y comprender, puedas utilizar tu conocimiento para ejecutar, implementar, calcular, resolver, o predecir.

Resultados de

Dominio de contenidos del temario

Las respuestas a la tareas de autoevaluación parecen indicar que domina los contenido de teoría y cálculo. Por otro lado, no domina tablas y figuras. Es recomendable que repase los contenidos de la asignatura, en especial tablas y figuras.

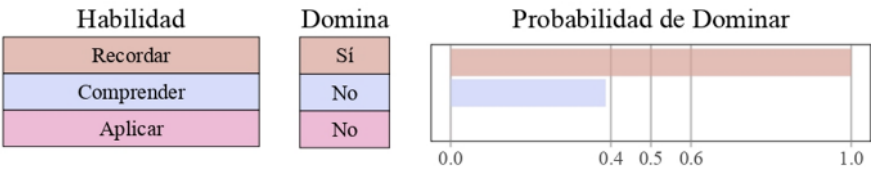
La siguiente figura indica las competencias dominadas, así como especifica la probabilidad de dominar cada una de ellas.



Dominio de habilidades

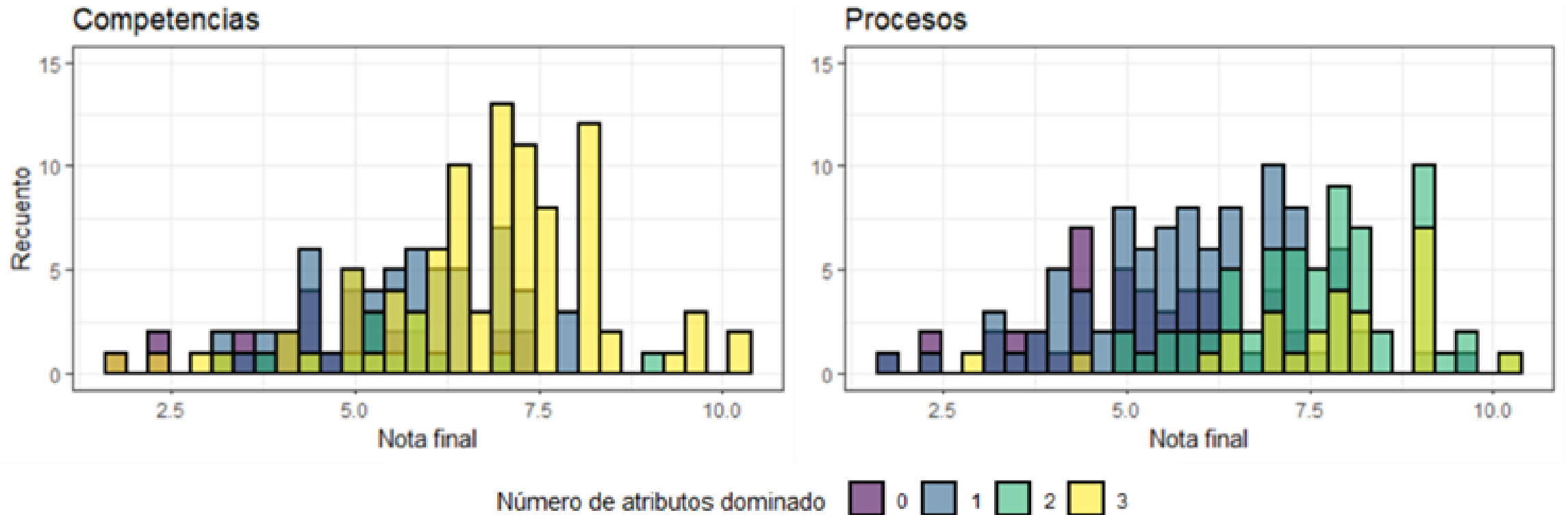
Las respuestas a la tareas de autoevaluación parecen indicar que de momento solamente es capaz de recordar con precisión los contenidos de la asignatura. Sin embargo, todavía encuentra problemas al comprender y aplicar el temario. Es recomendable que repase los contenidos de la asignatura y haga los ejercicios, en especial aquellos que exijan una mayor necesidad de comprender y aplicar los conocimientos.

La siguiente figura indica las habilidades dominadas, así como especifica la probabilidad de dominar cada una de ellas.



RELACIÓN CON LA NOTA FINAL:

- Las personas dominando al menos un atributo tenían entre 1.42 y 2.00 puntos en la nota final
- El tamaño del efecto fue mayor para **Comprensión (1.42)** y **Cálculo (1.15)**
- ...y menor para **Aplicar (0.86)** y **Figuras y tablas (0.90)**



Aplicación a una Asignatura

Conclusiones

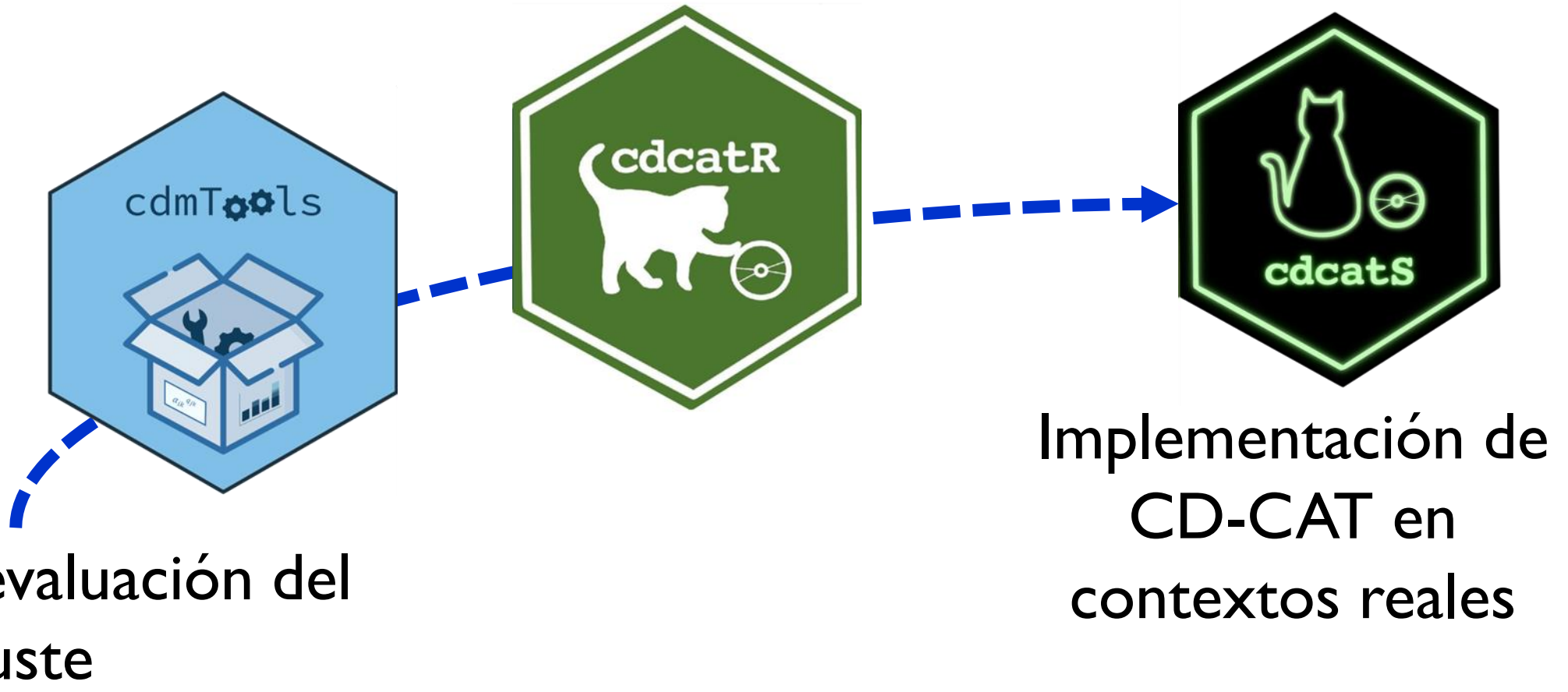
- Es posible extraer **información diagnóstica** relevante para la asignatura de Psicometría. **Extrapolable a otras asignaturas**
- El **Feedback grupal** (docentes) e **individual** (estudiantado) fue valorado positivamente
- Explorando otros formatos de ítems podrían **evaluarse procesos más complejos** (analizar, evaluar y crear)
- F💡Co es un **software** fácil de usar que posibilita estos análisis
- Los paquetes disponibles en R (`GDINA`, `cdmTools`) permitirían extraer más información
- Un test adaptativo informatizado aumentaría la eficiencia (`cdcatR`, `cdcatS`)

Notas finales

Notas Finales

Reglas de selección adicionales
(JSD, dual) y modelos

Futuros pasos:



Referencias seleccionadas

- de la Torre, J., & Sorrel, M. A. (2023). Cognitive Diagnosis Models. In F. Ashby, H. Colonius, & E. Dzhafarov (Eds.), *New Handbook of Mathematical Psychology* (pp. 385-420). Cambridge: Cambridge University Press.
<https://doi.org/10.1017/9781108902724.010>
- Sorrel, M. A., Olea, J., Abad, F. J., de la Torre, J., Aguado, D., & Lievens, F. (2016). Validity and reliability of situational judgement test scores: A new approach based on cognitive diagnosis models. *Organizational Research Methods*
- Shi, Q., Ma, W., Robitzsch, A., Sorrel, M. A., & Man, K. (2021). Cognitively diagnostic analysis using the G-DINA model in R. *Psych*
- Sanz, S., Kreitchmann, R. S., Nájera, P., Moreno, J. D., Martínez-Huertas, J. Á., & Sorrel, M. A. (2023). FoCo: A Shiny app for formative assessment using cognitive diagnosis modeling. *Educational Psychology*.
- Nájera, P., Abad, F. J., Chiu, C. Y., & Sorrel, M. A. (2023). The restricted DINA model: A comprehensive cognitive diagnostic model for classroom-level assessments. *Journal of Educational and Behavioral Statistics*.

*¡Muchas
gracias!*



Universidad Autónoma
de Madrid

Miguel A. Sorrel

@MigSorrel



/miguel-sorrel



miguel.sorrel@uam.es

