



UNIVERSITAT DE
BARCELONA

Facultat de Psicologia

DIAPPOSITIVES DE PSICOMETRIA

Curs 2017-2018

Autores:

Dra. Maite BARRIOS

Dra. M. Victoria CARRERAS

Dra. Juana GÓMEZ

Dra. Georgina GUILERA

Departament de Psicologia Social i Psicologia Quantitativa

Índex

1. Introducció a la psicometria.....	3
2. Anàlisi d'Ítems.....	13
3. Fiabilitat.....	36
4. Validesa.....	54
5. Baremació.....	91
6. Manual del Test.....	109
7. Annexes.....	114
Resum.....	115
Formulari	116

1. Introducció a la Psicometria

Lectura:

Muñiz, J. (2010). Las teorías de los tests: Teoría clásica y teoría de respuesta a los ítems. *Papeles del Psicólogo*, 31(1), 57-66.

Definició:

Etimològicament ➡ “Psyke” – “Metrum”

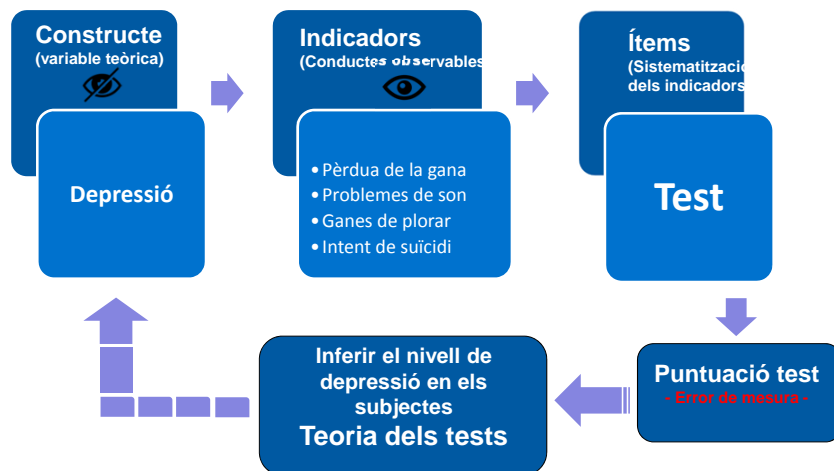
Conjunt de mètodes, tècniques i teories implicades en la mesura de les variables psicològiques. Tractaria de tot allò relacionat amb la mesura dels aspectes psicològics.

Especialització en aquelles propietats mètriques exigibles a les mesures psicològiques, independentment del camp substantiu d'aplicació i dels instruments emprats.

Com mesurem les característiques psicològiques?

- Les característiques psicològiques dels individus són inobservables, no es poden mesurar directament → **CONSTRUCTE**
- El constructe es pot manifestar a través d'una sèrie de conductes que sí són observables directament → **INDICADORS**
- Per avaluar el *nivell* dels subjectes en el *constructe* ens haurem de basar en els *indicadors*
- La sistematització d'un conjunt d'indicadors amb l'objectiu d'avaluar el nivell dels subjectes en un determinat constructe constitueix un **instrument de mesura o test**

Exemple: Depressió



CONTINGUTS

- **Teoria de la mesura:** fonamentació teòrica de la mesura

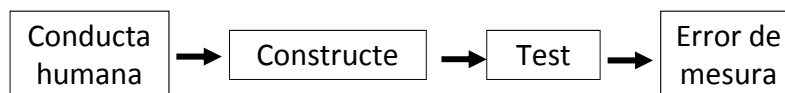
- **Escalament:** problemàtica de l'escalament dels estímuls psicològics i psicofísics

- **Teoria dels tests:** s'explicita la lògica i els models matemàtics subjacents a la construcció i ús dels tests
 - **Teoria clàssica dels tests**
 - **Teoria de resposta a l'ítem**
 - Tests referits al criteri

Instrument de mesura o test: Sistematització d'un **conjunt d'indicadors** amb la finalitat d'avaluar diversos nivells d'un constructe (variable teòrica inobservable que només es pot mesurar mitjançant indicadors observables que el representin).

Els **elements del test** constitueixen l'operativització del constructe ➡ han de reunir una sèrie de requisits per a demostrar que compleixen la funció per la qual han estat creats.

- Els instruments de mesura no són totalment precisos ➡ contenen **errors de mesura**
 - En quin grau les puntuacions del test indiquen el nivell de cada subjecte en el tret o característica mesurada?
 - Com ens podem assegurar que la puntuació obtinguda per un subjecte és la que realment li correspon?
 - En quin grau podem refiar-nos que, a partir de les puntuacions del test podem inferir el nivell del subjecte en el constructe mesurat?



Teoria dels Tests

- **Teoria dels tests** ➡ Funció matemàtica o model que estableix una relació entre:
 - variables observables (puntuacions en el test)
 - variables inobservables (puntuacions verdaderes o nivell dels subjectes en el tret mesurat pel test)

Teoria dels Tests

- Hi ha diferents models → tots ells analitzen la incidència dels errors en la mesura dels constructes psicològics i el grau en què afecten la fiabilitat dels instruments de mesura.
- Teoria dels tests més estesa → Teoria Clàssica dels Tests → model lineal.

Teoria Clàssica dels Tests (TCT)

Model lineal de la puntuació veritable proposat per Spearman:

$$X = V + e$$

On:

X – puntuació en el test

V – puntuació veritable, $E(X)$

e - error

Teoria Clàssica dels Tests (TCT)

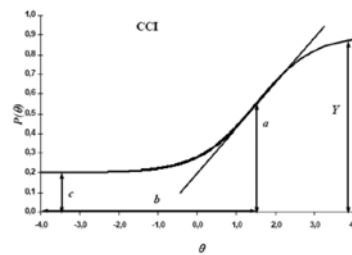
Mòdel	$X = V + e$
Supòsits	$V = E(X)$ Error aleatori: $E(e) = 0$ $r_{ve} = 0$ $r_{e_j e_k} = 0$
Test paral·lels	$V_j = V_k \quad i \quad \sigma_{e_j}^2 = \sigma_{e_k}^2$

Limitacions TCT

- Absència d'invariància respecte a l'instrument utilitzat.
- Absència d'invariància respecte a la mostra utilitzada per calcular les propietats dels tests.
- S'assumeix que la precisió del test és la mateixa, independentment del nivell del tret mesurat.
- No es disposa d'indicadors de bondat d'ajust que ens informin del grau en què el model s'ajusta a les dades.

Teoria de la Resposta a l'ítem (TRI)

- Corba Característica de l'ítem (CCI)
- La forma concreta de la CCI ve donada per tres paràmetres:
 - a: l'índex de discriminació
 - b: la dificultat de l'ítem
 - c: la probabilitat d'encertar l'ítem per l'atzar
- Supòsits:
 - Unidimensionalitat
 - Independència local



Models TRI

- Model logístic d'un paràmetre (*o model de Rasch*)

$$P(\theta) = \frac{1}{1 + e^{-(\theta - b_j)}}$$

- Model logístic de dos paràmetres

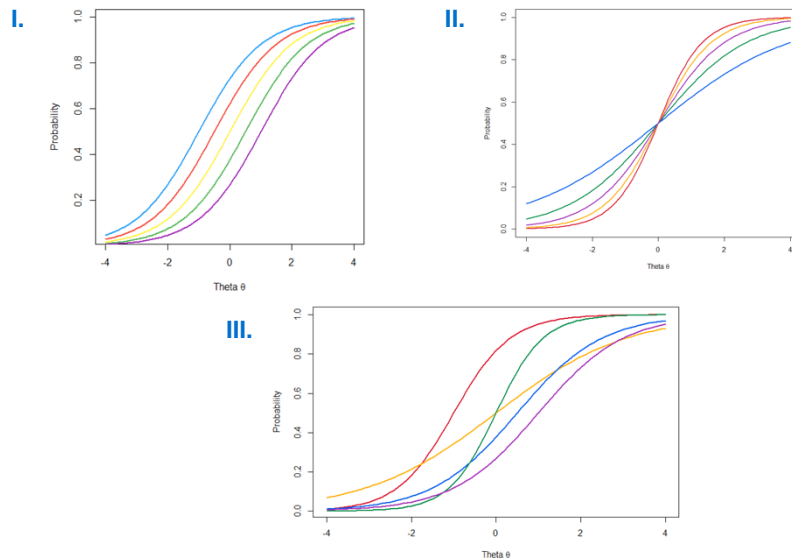
$$P(\theta) = \frac{1}{1 + e^{-Da_j(\theta - b_j)}}$$

- Model logístic de tres paràmetres

$$P(\theta) = c_j + (1 - c_j) \frac{1}{1 + e^{-Da_j(\theta - b_j)}}$$

On:
 $P(\theta)$: Probabilitat d'encertar l'ítem j si el nivell del tret és θ
 θ : nivell del tret o habilitat de la persona
 b_j : Paràmetre de dificultat de l'ítem j
 a_j : paràmetre de discriminació de l'ítem j
 e : base dels logaritmes neperians (2,718)
 D : Constant. $D=1,702$ (mètrica normal), $D=1$ (mètrica logística)
 c_j : probabilitat d'encert de l'ítem j per persones que tenen un nivell en el tret molt baix.

Exemples de CCI



TCT vs TRI

	TCT	TRI
Model	Lineal	No lineal
Interpretació	Puntuació total del test	Ítem
Precisió	La mateixa per a tots els participants / puntuacions	Estimada per cada participant / puntuació
Puntuació	Depèn dels ítems	Independent dels ítems
Paràmetres dels ítems	Depèn de la mostra utilitzada	És independent de la mostra utilitzada

Blocs del temari

- Anàlisi d'ítems
- Fiabilitat
- Validesa
- Baremació
- El manual del test

2. Anàlisi d'Ítems

Lectures:

Barbero García, M. I., Vila Abad, E. y Holgado Tello, F. P. (2010). *Psicometría*. Madrid: UNED. [pàgines: 72-80]

Abad, F. J., Olea, J., Ponsoda, V. y García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis. [pàgines: 38-46, 49-50, 54-57, 65-67]

Anàlisi d'ítems

- **Anàlisi qualitatiu:** judici d'experts sobre les característiques substancials, estilístiques, d'adequació i d'imparcialitat. Caldrà valorar:
 - Adequació del contingut
 - Format dels ítems per a la població diana i finalitat del test
 - Claredat en el redactat
 - Correcció gramatical
 - Seguiment de les regles bàsiques de redacció d'ítems

- **Anàlisi quantitatiu:** Càlcul d'indicadors per estudiar les característiques psicomètriques dels ítems a partir de les respostes obtingudes en mostres.
 - Índex de dificultat (ID)
 - Índex de discriminació (D)
 - Anàlisi dels distractors
 - Índex de fiabilitat
 - Índex de validesa

Directrius per la construcció dels ítems del test

- Definir l'**objectiu** de l'avaluació.
- **Context** en que s'utilitzaran els ítems (població objectiu, tipus d'administració i condicions en el que s'administrarà el test).
- **Domini** i context d'interès. Determinar el nivell del constructe i escollir els indicadors que el representen.

Elaboració dels ítems del test

Tancat o d'elecció

- Elecció binària
- Elecció múltiple
- Aparellament
- Format incomplet (*cloze*)
- Escales de classificació (*rating scales*)
- Llistats (*checklists*)

Obert o de construcció

- Resposta extensa
- Resposta curta

Recomanacions per la confecció d'ítems de resposta múltiple

- Cada ítem dirigit a avaluar un únic problema.
- Plausibilitat de les alternatives incorrectes.
- Ubicació a l'atzar de l'alternativa correcta.
- Evitar ítems que es puguin contestar de forma lògica o amb sentit comú.
- No repetir paraules o expressions per cada alternativa.
- Mateixa longitud de les alternatives.
- No utilitzar com alternativa: totes les anteriors o cap de les anteriors.
- Evitar les negacions dobles (en l'enunciat i en l'alternativa).

Errors més habituals en la confecció d'ítems

- No mateixa plausibilitat de les alternatives incorrectes.
- Enunciats en forma negativa.
- No mateixa longitud de les alternatives.
- Reiteracions innecessàries de text.
- Ús de l'alternativa: totes les anteriors o cap de les anteriors.

Anàlisi d'ítems: passos

- 1 • Decidir les principals propietats de les puntuacions del test.
- 2 • Identificar les anàlisis d'ítems més rellevants per a aquestes propietats.
- 3 • Administrar els ítems a una mostra representativa de la població a la qual està dirigit el test.
- 4 • Estimar les anàlisis identificades en el pas 2 per cada un dels ítems.
- 5 • Establir un pla per a la selecció d'ítems. Identificar i revisar els que funcionin malament.
- 6 • Seleccionar el bloc final d'ítems.

Anàlisi d'ítems: indicadors clàssics

- Índex de dificultat (ID)
- Índex de discriminació (D)
- Anàlisi dels distractors
- Índex de fiabilitat de l'ítem
- Índex de validesa de l'ítem

Anàlisi d'ítems: Índex de Dificultat

ID sense corregir per atzar: $ID = \frac{A}{N}$

ID corregit per atzar: $ID = \frac{A - E/(k-1)}{N}$

Punt. corregida del subjecte (Pc): $Pc = A - E/(k-1)$

Mitjana del test = Suma ID dels ítems

$$\bar{x} = \frac{\sum x_j}{N} = \sum ID_j$$

Exemple: Índex Dificultat sense corregir per atzar

Subj.	Ítems						Total
	1	2	3	4	5	6	
1	0	1	0	1	0	1	3
2	1	1	1	0	0	1	4
3	0	1	0	1	1	1	4
4	1	1	0	1	0	1	4
5	1	1	0	1	0	0	3
6	0	1	0	0	1	0	2
7	1	1	1	1	1	1	6
8	0	0	0	0	1	0	1
Total A	4	7	2	5	4	5	
ID	0.5	0.875	0.25	0.625	0.5	0.625	

$$\bar{x} = \frac{\sum x_j}{N} = \frac{27}{8} = 3.38$$

$$\sum ID_j = 3.38$$

Exercici: Índex Dificultat sense corregir per atzar

Subj.	Ítems							
	1	2	3	4	5	6	7	8
1	0	0	0	1	1	0	1	1
2	1	0	0	1	0	0	1	1
3	1	1	0	1	1	0	1	1
4	1	0	1	1	1	1	1	0
5	0	1	0	1	1	1	1	0
6	1	1	1	1	0	0	1	1
7	1	0	0	0	0	1	1	0
8	1	0	0	1	1	0	1	1
9	1	1	1	1	1	1	1	1
10	0	1	0	1	0	0	1	1

Comproveu que la mitjana del test equival a la suma dels ID

Exemple: Índex Dificultat corregit per atzar

Subj.	Ítems (5 alternatives de resposta)								Totals			
	1	2	3	4	5	6	7	8	A	E	NC	Pc
1	1	1	0	1	X	1	1	X	5	1	2	4.75
2	0	1	0	1	1	0	X	1	4	3	1	3.25
3	1	1	1	1	X	1	1	0	6	1	1	5.75
4	1	0	0	1	1	0	X	0	3	4	1	2
5	1	1	1	0	1	1	1	1	7	1	0	6.75
6	1	0	1	1	1	X	0	0	4	3	1	3.25
7	1	1	1	1	1	X	1	0	6	1	1	5.75
8	0	0	1	1	1	1	X	1	5	2	1	4.5
9	1	0	1	1	X	1	X	1	5	1	2	4.75
10	1	X	1	X	1	1	1	1	6	0	2	6
A	8	5	7	8	7	6	5	5	1 – encert			
E	2	4	3	1	0	2	1	4	0 – error			
Cont.	10	9	10	9	7	8	6	9	X – no contestat			

$\sum Pc = 46.75$

Exemple: Índex Dificultat corregit per atzar

$$ID_1 = \frac{8 - 2/4}{10} = 0.75$$

$$ID_5 = \frac{7 - 0/4}{10} = 0.7$$

$$ID_2 = \frac{5 - 4/4}{10} = 0.4$$

$$ID_6 = \frac{6 - 2/4}{10} = 0.55$$

$$ID_3 = \frac{7 - 3/4}{10} = 0.625$$

$$ID_7 = \frac{5 - 1/4}{10} = 0.475$$

$$ID_4 = \frac{8 - 1/4}{10} = 0.775$$

$$ID_8 = \frac{5 - 4/4}{10} = 0.4$$

$$\bar{x} = \frac{46.75}{10} = 4.68$$

$$\sum ID = 4.68$$

Exercici: Índex Dificultat corregit per atzar

Subj	Ítems (5 alternatives de resposta)							
	1	2	3	4	5	6	7	8
1	1	1	X	1	X	0	0	1
2	1	1	1	1	1	X	1	1
3	0	1	1	0	1	1	1	1
4	0	1	1	0	0	0	0	1
5	0	0	0	1	0	0	0	1
6	1	1	0	1	0	0	0	0
7	1	1	0	X	0	1	1	1
8	X	1	1	X	1	X	1	1
9	1	0	1	1	1	X	0	1
10	1	0	1	0	0	1	0	1

Calculeu els ID dels ítems i les Pc dels subjectes

1 - Encerts

0 - Errors

X - No contesta

Anàlisi d'ítems: Índex de Discriminació

Avalua si els ítems són capaços de posar de manifest les diferències individuals entre els subjectes en la variable que mesurem



Un ítem té poder discriminant si els subjectes que tenen un nivell alt en la variable encerten més l'ítem o obtenen puntuacions més elevades que els subjectes que tenen un nivell baix en la variable

Anàlisi d'ítems: Índex de Discriminació

Posa en relació l'execució del subjecte en l'ítem amb la seva execució global en el test



Càlcul:

- Índex basat en la diferència de les proporcions d'encert o de respostes positives a l'ítem
- Índexs basats en les correlacions ítem-test

Criteris per a la valoració de la discriminació de l'ítem (Ebel & Frisbie, 1991)

Índex de discriminació	Valoració de l'ítem
≥ 0,40	Ítem amb molt bona capacitat discriminativa
0,30 – 0,39	Ítem amb una capacitat discriminativa raonablement bona, però susceptible a ser millorat
0,20 – 0,29	Límit. Es requereix millorar l'ítem
< 0,20	Cal eliminar l'ítem o modificar-lo completament

Discriminació: índex basat en la diferència de les proporcions d'encert o respostes positives

$$D = \frac{n_s}{N_s} - \frac{n_i}{N_i}$$

On:

n_s – nombre persones que encerten ítem o donen una resposta positiva en el grup superior

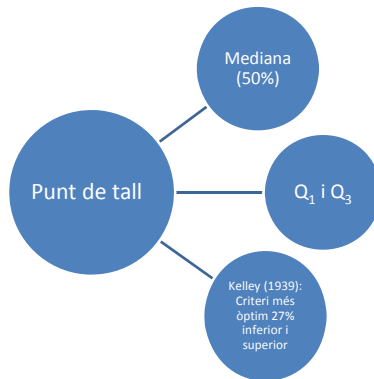
n_i – nombre persones que encerten ítem o donen una resposta positiva en el grup inferior

N_s – nombre persones que responen en el grup superior

N_i – nombre persones que responen en el grup inferior

Discriminació: índex basat en la diferència de les proporcions d'encert o respostes positives

Rang: $-1 \leq D \leq 1$, $D=1$: màxima capacitat discriminativa



Discriminació: exemple diferències

Grups basats en la mediana (50%)

Subj.	Ítems						Punt. Total	Grup
	1	2	3	4	5	6		
1	0	1	0	1	0	1	3	Inf
2	1	1	1	0	0	1	4	Sup
3	0	1	0	1	1	1	4	Sup
4	1	1	0	1	0	1	4	Sup
5	1	1	0	1	0	0	3	Inf
6	0	1	0	0	1	0	2	Inf
7	1	1	1	1	1	1	6	Sup
8	0	0	0	0	1	0	1	Inf
n_s	3	4	2	3	2	4		
n_i	1	3	0	2	2	1		
D	0.5	0.25	0.5	0.25	0	0.75		

Exercici: Índex de discriminació basat en la diferència de les proporcions d'encert o respostes positives

Índex de discriminació de l'ítem 1

Subjecte	ítem 1	ítem 2	ítem 3	ítem 4	ítem 5	ítem 6	ítem 7	ítem 8	Total
1	0	0	1	1	1	1	0	1	5
2	1	1	1	1	1	1	1	1	8
3	0	0	0	1	0	1	0	1	3
4	0	1	0	0	0	0	0	0	1
5	0	1	1	0	0	1	0	0	3
6	1	1	1	1	1	1	1	1	8
7	1	0	0	1	1	0	1	0	4
8	1	1	1	0	0	1	0	0	4
9	0	1	1	1	1	1	1	1	7
10	1	0	0	1	1	1	0	1	5

Discriminació: Índexs basats en la correlació ítem-test

Rang: $-1 \leq D \leq 1$, $D=1$: màxima capacitat discriminativa

Típus de Correlacions		
Ítem	Test	Correlació
Quantitativa	Quantitativa	Pearson r
Dicotòmica	Quantitativa	Biserial-Puntual r_{bp}
Dicotomitzada	Quantitativa	Biserial r_b
Dicotòmica	Dicotòmica	Fi ϕ
Dicotomitzada	Dicotomitzada	Tetracòrica r_t

Restar la puntuació de l'ítem en la del test o
utilitzar la fórmula de correcció de la correlació

Coefficient de Correlació de Pearson

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y}$$

On:

- S_{xy} Covariància test-ítem
- S_x desviació típica del test
- S_y desviació típica de l'ítem

Coefficient de Correlació Biserial-Puntual

$$r_{bp} = \frac{\bar{x}_p - \bar{x}_x}{S_x} \sqrt{\frac{p}{q}}$$

On:

- \bar{x}_p mitjana en el test dels subjectes que encerten o donen una resposta positiva a l'ítem
- \bar{x}_x mitjana del test
- S_x desviació típica del test
- p proporció de subjectes que encerten o donen una resposta positiva a l'ítem
- q (1 - p)

Coefficient de Correlació Biserial

$$r_b = \frac{\bar{x}_p - \bar{x}_x}{S_x} \frac{p}{y}$$

On:

- \bar{x}_p mitjana en el test dels subjectes que encerten o donen una resposta positiva a l'ítem
- \bar{x}_x mitjana del test
- S_x desviació típica del test
- p proporció de subjectes que encerten o donen una resposta positiva a l'ítem
- y ordenada que correspon en la distribució normal al valor que deixa per sota una àrea igual a p

Coefficient Fi ϕ

$$\Phi = \frac{p_{xy} - p_x p_y}{\sqrt{p_x q_x p_y q_y}}$$

On:

- p_{xy} proporció d'encerts o respostes positives en x i y
- p_x proporció d'encerts o respostes positives en x
- p_y proporció d'encerts o respostes positives en y
- q_x 1 - p_x
- q_y 1 - p_y

Coefficient Tetracòric

		Test		
		0	1	
ítem	0	a	b	p
	1	c	d	q
		p'	q'	n

$$r_t = \sqrt{\frac{(ad) - (bc)}{n^2 y y'}}$$

On:

- y ordenada de la corba normal que separa les dues categories del test (grup superior/inferior)
- y' ordenada de la corba normal que separa les dues categories de l'ítem (encert/error)
- a, b, c i d fi en la taula de contingència 2 x 2 entre el test i l'ítem

Exemple: Correlació biserial-puntual

Subj.	Ítems						Totals	
	1	2	3	4	5	6	Total	Total-4
1	0	1	0	1	0	1	3	2
2	1	1	1	0	0	1	4	4
3	0	1	0	1	1	1	4	3
4	1	1	0	1	0	1	4	3
5	1	1	0	1	0	0	3	2
6	0	1	0	0	1	0	2	2
7	1	1	1	1	1	1	6	5
8	0	0	0	0	1	0	1	1
Total	4	7	2	5	4	5		

r_{bp} de l'ítem 4?

$$p = \frac{5}{8} = 0.625$$

$$q = \frac{3}{8} = 0.375$$

$$\bar{x} = \frac{2+4+3+3+2+2+5+1}{8} = 2.75$$

$$\bar{x}_p = \frac{2+3+3+2+5}{5} = 3$$

$$S_x^2 = \frac{2^2+4^2+3^2+3^2+2^2+2^2+5^2+1^2}{8} - 2.75^2 = 1.44 \Rightarrow S_x = 1.20$$

$$r_{bp} = \frac{3 - 2.75}{1.2} \sqrt{\frac{0.625}{0.375}} = 0.27$$

Exemple: Correlació biserial-puntual. Ítem 5

Item5	Total-5
0	3
0	4
1	3
0	4
0	3
1	1
1	5
1	0

$$p = 0.5$$

$$q = 0.5$$

$$\bar{x}_p = \frac{3+1+5+0}{4} = 2.25 \quad \bar{x}_x = 2.88$$

$$S_x^2 = \frac{85}{8} - 2.88^2 = 2.36 \rightarrow S_x = 1.54$$

$$r_{bp} = \frac{2.25 - 2.88}{1.54} \sqrt{\frac{0.5}{0.5}} = -0.41$$

Exercici: Correlació biserial-puntual. Ítem 6

Subj.	Ítems						Total
	1	2	3	4	5	6	
1	0	1	0	1	0	1	3
2	1	1	1	0	0	1	4
3	0	1	0	1	1	1	4
4	1	1	0	1	0	1	4
5	1	1	0	1	0	0	3
6	0	1	0	0	1	0	2
7	1	1	1	1	1	1	6
8	0	0	0	0	1	0	1
Total	4	7	2	5	4	5	

Correcció de la correlació

En cas de no haver eliminat l'ítem per calcular la correlació, caldrà aplicar la correcció següent:

$$Cr = \frac{r_{jx} S_x - S_j}{\sqrt{S_x^2 + S_j^2 - 2r_{jx} S_x S_j}}$$

On:

r_{jx} Coeficient de correlació entre l'ítem i el test

S_x Desviació típica del test

S_j Desviació típica de l'ítem

S_x^2 Variància del test

S_j^2 Variància de l'ítem

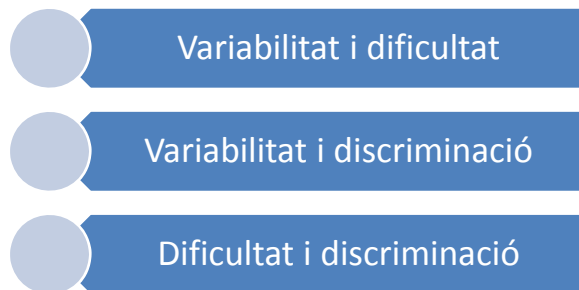
Nota: Cal recordar que la variància per variables dicotòmiques:

$$S_j^2 = p \cdot q$$

Exercici: Correlació biserial-puntual. Ítem 6 (sense restar puntuació de l'ítem – aplicant la correcció de la correlació)

Item 6	Total
1	3
1	4
1	4
1	4
0	3
0	2
1	6
0	1

Relacions entre indicadors



Variabilitat i dificultat

Què ocorre si tothom encerta un ítem?
I si tothom el falla?



$$S_j^2 = 0$$

L'ítem no és útil per establir diferències entre persones

Variabilitat i dificultat

En ítems dicotòmics (1 = encert i 0 = error)...

$$S_j^2 = p \cdot q$$

On:

p proporció d'encerts

q proporció d'errors (1-p)

S_j^2 màxima quan $p = 0,5$ i $q = 0,5 \rightarrow ID_j = 0,5$

S_x^2 s'aproxima al valor màxim quan els ID dels ítems que
composen el test són més propers a 0,5

Variabilitat i discriminació

Ítem pot discriminar si $S_j^2 > 0$

Mesura del poder discriminatiu d'un test $\rightarrow S_x^2$

$$S_x = \sum S_j \cdot r_{jx}$$

On:

S_x desviació típica del test

S_j desviació típica de l'ítem j

r_{jx} índex de discriminació (correlació ítem-test)

Dificultat i discriminació

Els valors que pot prendre l'índex de discriminació d'un ítem depèn del valor de l'índex de dificultat.

Ítem fàcil \rightarrow permet discriminar entre les persones amb nivells baixos en el tret mesurat (puntuacions baixes)

Ítem difícil \rightarrow permet discriminar entre les persones amb nivells elevats en el tret mesurat (puntuacions altes)

Dificultat i discriminació

Major capacitat discriminativa del test quan conté:



- Majoria d'ítems de dificultat mitjana (en ítems dicotòmics ID al voltant de 0,5)
- Alguns ítems fàcils per discriminar entre persones amb menys nivell
- Alguns ítems difícils per discriminar entre persones amb més nivell

Anàlisi d'ítems: Anàlisi dels distractors Respostes d'elecció múltiple

Els distractors han de ser igualment atractius per als subjectes que no coneguin la resposta



Tenir la mateixa probabilitat de ser escollits a l'atzar quan el subjecte intenta endevinar la resposta correcta.

Exemple: Anàlisi dels distractors

		Alternatives					Total
		A	B	C	D	E	
Ítem	50% superior	5	7	47	7	4	70
	50% inferior	11	9	32	8	10	70
Percentatge		11.43%	11.43%	56.43%	10.71%	10%	

		Alternatives					Total
		A	B	C	D	E	
Ítem	50% superior	0	20	9	8	33	70
	50% inferior	2	15	14	18	21	70
Percentatge		1.43%	25%	16.43%	18.57%	38.57%	

Exercici: Anàlisi dels distractors

		Alternatives					Total
		A	B	C	D	E	
Ítem	50% superior	6	93	6	41	4	150
	50% inferior	7	59	37	5	42	150

Anàlisi dels distractors a partir de D

Exemple:

	Alternatives				Total
	A	B	C	D	
50% superior	14	55	18	13	100
50% inferior	20	35	26	19	100

Índex de Discriminació de l'alternativa A:
$$D = \frac{n_s}{N_s} - \frac{n_i}{N_i} = \frac{14}{100} - \frac{20}{100} = -0,06$$

Índex de Discriminació de l'alternativa C:
$$D = \frac{n_s}{N_s} - \frac{n_i}{N_i} = \frac{18}{100} - \frac{26}{100} = -0,08$$

Índex de Discriminació de l'alternativa D:
$$D = \frac{n_s}{N_s} - \frac{n_i}{N_i} = \frac{13}{100} - \frac{19}{100} = -0,06$$

Anàlisi dels distractors a partir de D

Exercici:

	Alternatives				Total
	A	B	C	D	
50% Grup superior	12	3	40	45	100
50% Grup inferior	27	20	25	28	100

Índex de Fiabilitat de l'ítem

Estimació de la fiabilitat amb la que l'ítem mesura la característica que mesura el test en conjunt

$$IF = S_j D_j$$

On:

S_j desviació típica de les puntuacions en l'ítem j

D_j índex de discriminació de l'ítem j

Índex de Validesa de l'ítem

- Contribució de l'ítem en la predicció d'una variable externa (criteri)
- Correlació de l'ítem amb un criteri extern
- El tipus de correlació està en funció del tipus d'ítem i el tipus de criteri

<i>Tipus de Correlacions</i>		
<i>Ítem</i>	<i>Criteri</i>	<i>Correlació</i>
Quantitativa	Quantitativa	Pearson r
Dicotòmica	Quantitativa	Biserial-Puntual r_{bp}
Dicotomitzada	Quantitativa	Biserial r_b
Dicotòmica	Dicotòmica	Fi ϕ
Dicotomitzada	Dicotomitzada	Tetracòrica r_t

Exemple: Índex de Validesa de l'ítem

Ítem dicotòmic d'un qüestionari de depressió (0=no depressió, 1=depressió)
Criteri extern: puntuació total d'un test d'ansietat

Ítem depressió	Test d'ansietat
0	3
1	12
1	10
0	2
1	15
0	4
0	3
1	11
1	15
1	16
1	18
0	0

$$p = \frac{7}{12} = 0.583 \quad q = \frac{5}{12} = 0.417$$

$$\bar{x}_p = \frac{12+10+15+11+15+16+18}{7} = 13.86 \quad \bar{x}_x = \frac{109}{12} = 9.08$$

$$S_x^2 = \frac{3^2+12^2+10^2+2^2+15^2+4^2+3^2+11^2+15^2+16^2+18^2+0^2}{12} - 9.08^2 = 36.97$$

$$S_x = 6.08$$

$$r_{bp} = \frac{13.86 - 9.08}{6.08} \sqrt{\frac{0.583}{0.416}} = 0.93$$

3. Fiabilitat

TEMA: Fiabilitat

Lectura:

Prieto, G., y Delgado, A.R. (2010). Fiabilidad y validez. *Papeles del Psicólogo*, 31(1), 67-74.

Definició

Fiabilitat: Absència d'error de mesura (consistència o estabilitat de les mesures quan el procés de mesura es repeteix)

$$X = V + e$$

$$r_{xx} = \frac{S_V^2}{S_x^2}$$

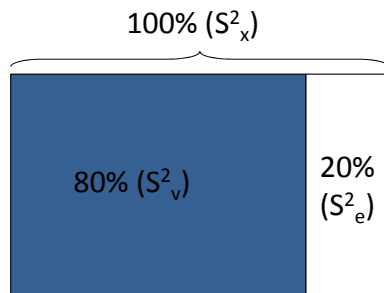
On:

S_V^2 - variància vertadera del test

S_x^2 - variància obtinguda del test

Què indica un coeficient de fiabilitat?

$$r_{xx} = 0.80$$



Del 100% de la variància obtinguda (S^2_x), el 80% és deguda a la variància vertadera (S^2_v) i per tant, el 20% restant és deguda a la variància d'error (S^2_e)

Procediments per estimar la fiabilitat

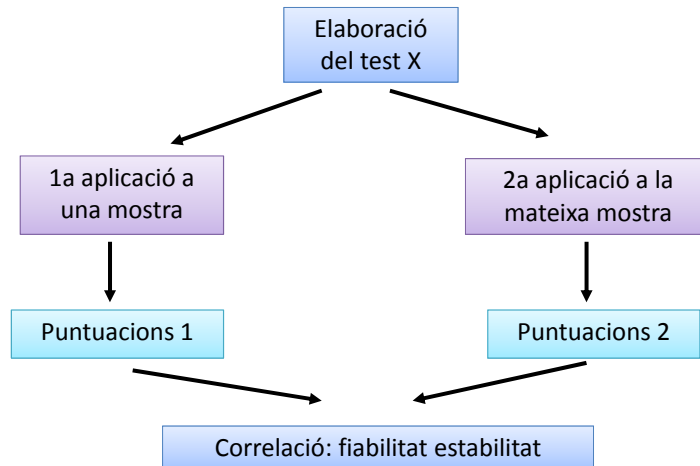
Estabilitat de la mesura → Test-retest

Equivalència de les mesures → Formes paral·leles

Consistència interna → Intercorrelacions entre els ítems:

- ❖ Dues meitats - Correcció Spearman-Brown
- ❖ Coeficient alfa de Cronbach (anàlisi de les covariàncies entre els ítems)

Fiabilitat test-retest



Críteris valoració fiabilitat test-retest (estabilitat mesura)

Prieto, G. y Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 77, 65-75.

Inadeuada: $r < 0.55$

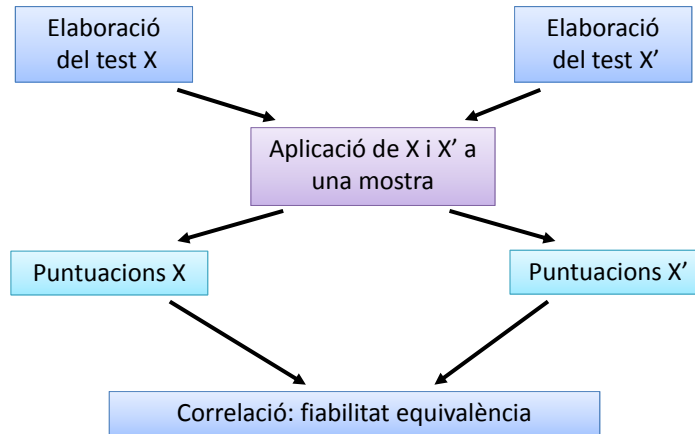
Adeuada però amb algunes mancances: $0.55 \leq r < 0.65$

Adeuada: $0.65 \leq r < 0.75$

Bona: $0.75 \leq r < 0.80$

Excel·lent: $r \geq 0.80$

Fiabilitat formes paralel·les



Críteris valoració fiabilitat formes paralel·les (equivalència mesura)

Prieto, G. y Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 77, 65-75.

Inadeuada: $r < 0.50$

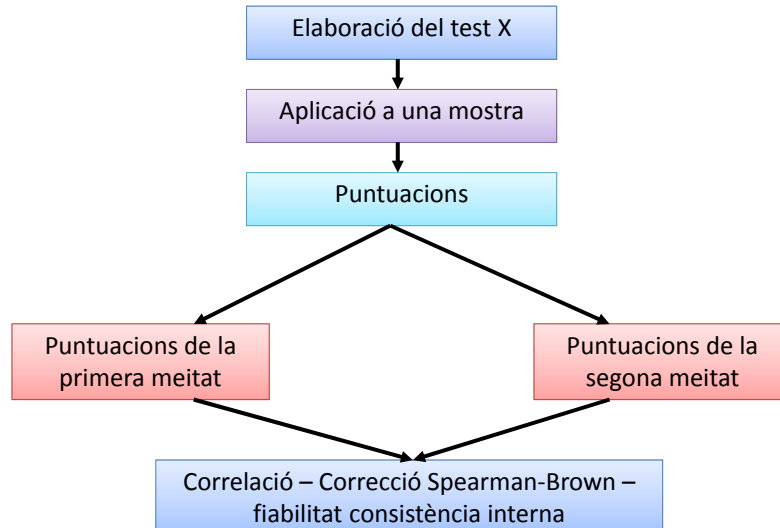
Adeuada però amb algunes mancances: $0.50 \leq r < 0.60$

Adeuada: $0.60 \leq r < 0.70$

Bona: $0.70 \leq r < 0.80$

Excel·lent: $r \geq 0.80$

Fiabilitat dues meitats



Quins ítems han de formar part de cada una de les dues meitats?

Ítems meitat A	Ítems meitat B	Correlació entre les dues meitats
1, 2, 3	4, 5, 6	0.559
1, 2, 4	3, 5, 6	0.645
1, 2, 5	3, 4, 6	0.474
1, 2, 6	3, 4, 5	0.586
1, 3, 4	2, 5, 6	0.653
1, 4, 5	2, 3, 6	0.730
1, 5, 6	2, 3, 4	0.507
2, 3, 5	1, 4, 6	0.426
2, 4, 5	1, 3, 6	0.824
2, 4, 6	1, 3, 5	0.546

Extret de Viladrich i cols.
(2009)



Correcció d'Spearman-Brown pel càlcul de les dues meitats

$$r_{xx} = \frac{2 r_{x_1 x_2}}{1 + r_{x_1 x_2}}$$

On:

r_{xx} - Fiabilitat del test sencer

$r_{x_1 x_2}$ - Correlació del subtest 1 i el subtest 2

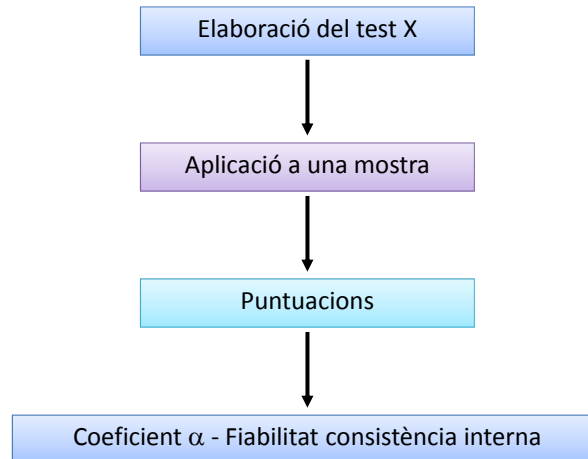
Exemple Correcció d'Spearman-Brown pel càlcul de les dues meitats

Aplicació de la correcció d'Spearman-Brown a les dades de l'exemple anterior:

$$r_{x_1 x_2} = 0,546$$

$$r_{xx} = \frac{2 \cdot 0,546}{1 + 0,546} = 0,7063$$

Fiabilitat Alfa de Cronbach



Coeficient alfa (Cronbach, 1951)

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum S_j^2}{S_x^2} \right)$$

On:

n nombre d'ítems del test

$\sum S_j^2$ suma de les variàncies dels n ítems

S_x^2 variància de la puntuació total del test

Variacions del coeficient alfa: Kuder-Richardson (1937)

Aquest és el coeficient alfa per a ítems dicotòmics.

$$\alpha = KR20 = \frac{n}{n-1} \left(1 - \frac{\sum p_j \cdot q_j}{S_x^2} \right)$$

On:

n nombre d'ítems del test
p_j proporció de respostes positives en l'ítem
q_j proporció de respostes negatives en l'ítem
S_x² variància de la puntuació total del test

Criteris valoració fiabilitat consistència interna

Prieto, G. y Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 77, 65-75.

Inadeuada: $r < 0.60$

Adeuada però amb algunes mancances: $0.60 \leq r < 0.70$

Adeuada: $0.70 \leq r < 0.80$

Bona: $0.80 \leq r < 0.85$

Excel·lent: $r \geq 0.85$

Exemple coeficient alfa

¿Quin serà el valor del coeficient alfa si
 $n = 40$, $\sum S_j^2 = 2$ i $S_x^2 = 6$?

$$\alpha = \frac{40}{40 - 1} \left(1 - \frac{2}{6} \right) = 0.68$$

Contribució dels ítems a la consistència interna

- ❖ Cada ítem que forma part d'un instrument de mesura té una contribució particular a la consistència interna d'aquest
- ❖ Podem calcular el coeficient α per al conjunt d'ítems del test però havent eliminat prèviament del càlcul les dades de l'ítem que volem avaluar

$$\alpha = 0.72$$

	Alfa sense l'ítem
Item 1	0.71
Item 2	0.68
Item 3	0.70
Item 4	0.76
Item 5	0.71



Factors que influeixen en la fiabilitat del test

- ❖ Mostratge
- ❖ Variabilitat
- ❖ Longitud del test (nombre d'ítems)
- ❖ Nivell de les puntuacions en el test

Formes d'incrementar la fiabilitat del test

- ❖ Augmentar el nombre d'ítems - Fórmula d'Spearman-Brown
- ❖ Eliminar ítems problemàtics (baixes correlacions ítem-test)
- ❖ Millorar les condicions d'aplicació

Relació entre fiabilitat i longitud del test Fórmula d'Spearman-Brown

$$R_{xx} = \frac{k r_{xx}}{1 + (k - 1) r_{xx}}$$

On:

R_{xx} fiabilitat del test allargat

r_{xx} fiabilitat del test abans d'allargar-lo

k nombre de vegades que hem allargat el test:

$$k = \frac{n_i + n_a}{n_i}$$

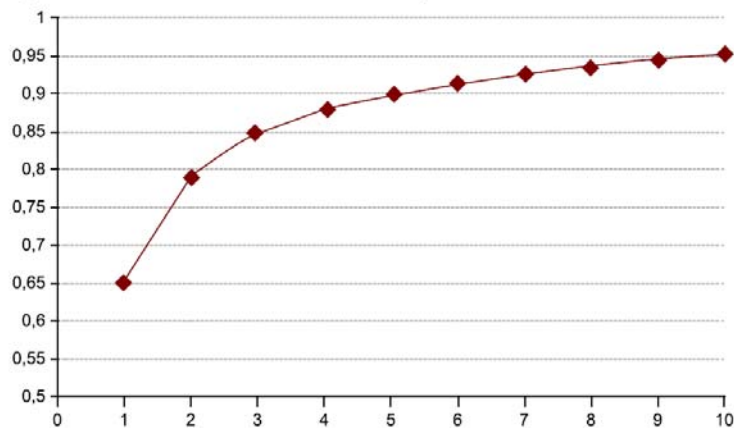
On:

n_i - nombre d'ítems inicials

n_a - nombre d'ítems afegits

Relació entre fiabilitat i longitud del test

Figura 1. Relación entre el coeficiente de fiabilidad y el aumento de ítems en un test



Extret de: Meneses, J., Barrios, M., Bonillo, A., Coscolluela, A., Lozano, L. M., Turbany, J., Valero, S. (2013). Psicometría. Barcelona: Editorial UOC.

Relació entre fiabilitat i longitud del test

Suposem que la fiabilitat d'un test amb 20 ítems és de 0.70. ¿Quina seria la seva fiabilitat si afegim 40 ítems al test?

$$k = \frac{20 + 40}{20} = 3$$

$$R_{xx} = \frac{3 \cdot 0.7}{1 + (3 - 1) 0.7} = 0.875$$

Longitud del test per aconseguir una fiabilitat determinada

$$k = \frac{R_{xx} (1 - r_{xx})}{r_{xx} (1 - R_{xx})}$$

On:

R_{xx} fiabilitat que es pretén aconseguir

r_{xx} fiabilitat del test original

k nombre de vegades que hem allargat el test:

$$k = \frac{n_i + n_a}{n_i}$$

On:

n_i - Nombre d'ítems inicials

n_a - Nombre d'ítems afegits

Longitud del test per aconseguir una fiabilitat determinada

Suposem que un test amb 10 ítems té una fiabilitat de 0.60. ¿Quants ítems caldrà afegir per aconseguir una fiabilitat de 0.75?

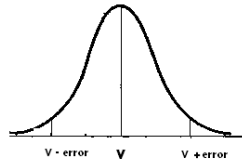
$$k = \frac{0.75 (1 - 0.60)}{0.60 (1 - 0.75)} = 2$$

$$2 = \frac{10 + n_a}{10} \rightarrow n_a = 2 \cdot 10 - 10 = 10$$

Exercici: Longitud del test per aconseguir una fiabilitat determinada

Suposem que un test amb 50 ítems té una fiabilitat de 0.75. ¿Quants ítems caldrà afegir per aconseguir una fiabilitat de 0.90?

Estimació per interval de la puntuació vertadera - Error típic de mesura



$$V_{\text{int}} = x_i \pm z \cdot S_e$$

On:

x_i - Puntuació obtinguda pel subjecte en el test

S_e - Error típic de mesura:

$$S_e = S_x \sqrt{1 - r_{xx}}$$

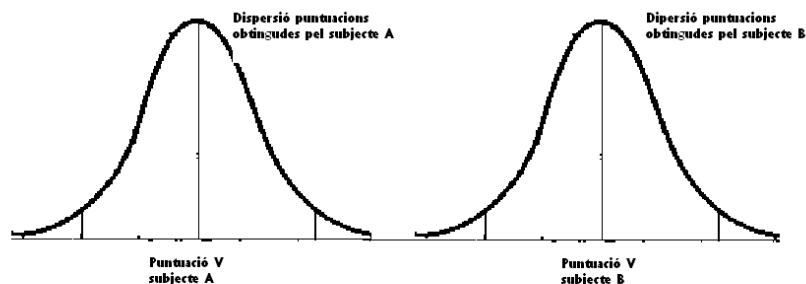
On:

S_x - desviació típica del test

r_{xx} - coeficient fiabilitat del test

27

Error típic de mesura - TCT



Desviació típica dels errors la mateixa o molt semblant per totes les persones $\rightarrow S_e$

28

Aplicacions de l'error típic de mesura - TCT

1. Estimació per interval de la puntuació vertadera
2. Comparació de les puntuacions observades de dues persones
3. Valoració del canvi de la puntuació observada d'una persona després d'una intervenció

29

Exemple: Estimació per interval de la puntuació vertadera - Error típic de mesura

Suposem que el coeficient de fiabilitat d'un test és de 0.75, la seva mitjana 80 i la desviació típica 16. ¿Quin és l'error típic de mesura? ¿En quin interval se situa la puntuació vertadera d'un subjecte que té una puntuació observada de 86 a un nivell de confiança del 95% ($z=1.96$)?

$$S_e = 16 \sqrt{1 - 0.75} = 8$$

$$V_{\text{int}} = 86 \pm 1.96 \cdot 8 \rightarrow 70.32 \leq V \leq 101.68$$

30

Exercici: Estimació per interval de la puntuació vertadera - Error típic de mesura

Suposem que el coeficient de fiabilitat d'un test és de 0.80, la seva mitjana 15 i la variància 9. ¿Quin és l'error típic de mesura? ¿En quin interval se situa la puntuació vertadera d'un subjecte que té una puntuació observada de 12, a un nivell de confiança del 95% ($z=1.96$)?

Exercici: Comparar puntuacions observades de dues persones

En l'escala d'Extraversió d'una prova dues persones obtenen unes puntuacions observades de 12 i 14 punts, respectivament. Sabem que la desviació típica del test és de 2,7 i el coeficient de fiabilitat de 0.80.

Fins a quin punt podem concloure que ambdues persones difereixen en el seu nivell d'extraversió (NC = 95%; $z = 1.96$)?

Exercici 1: Fiabilitat

Un test de 32 ítems s'aplica a 200 subjectes i s'obté una mitjana de 23 i una desviació típica de 4. Al dividir el test en ítems parells i imparells, la correlació entre les dues meitats és de 0,91.

- 1) Quin és el coeficient de fiabilitat del test? Interpreta aquest coeficient i comenta si el valor és o no acceptable.
- 2) Quants ítems caldria afegir a aquest test per a elevar la seva fiabilitat fins a 0,99? Val la pena modificar el test per a augmentar la fiabilitat fins a aquest nivell?
- 3) Quant val el coeficient alfa d'aquest test, tenint en compte que la suma de les variàncies dels seus ítems és 1,25? Interpreta aquest coeficient i comenta si el valor és acceptable.
- 4) En el test original un subjecte ha arribat a una puntuació de 26. Entre quines puntuacions s'espera que es trobi la puntuació veritable del subjecte al nivell de confiança del 95%? Interpreta l'interval obtingut.

Exercici 2: Fiabilitat

Es va aplicar un qüestionari per identificar factors de risc interpersonal en el consum de drogues en adolescents a una mostra de 2126 estudiants d'ESO. El test està format per 51 ítems que s'agrupen en 7 factors. A continuació es presenta la consistència interna per a cadascun dels factors.

Factors	α
1. Reacció familiars	0,887
2. Amics	0,860
3. Accessibilitat	0,891
4. Situacions de risc	0,640
5. cació en drogues	0,850
6. Estil educatiu	0,740
7. Activitats protectores	0,706

- a) Comenta els valors de fiabilitat dels diferents factors.
- b) Sabent que la suma de las variàncies dels ítems és 1,45, la desviació típica del test és 4,25, calcula l'alfa de Cronbach del test.
- c) Quina seria la fiabilitat del test en el cas de què el qüestionari estigués format per 60 ítems?
- d) Si volem aconseguir que el qüestionari tingui una fiabilitat de 0,98, quants ítems caldria afegir?
- e) Suposem que un dels adolescents que va contestar el qüestionari va obtenir una puntuació directa de 73. Coneixent les dades de l'apartat b), en quin interval es trobaria la seva puntuació veritable amb un nivell de confiança del 95%?

4. Validesa

TEMA: Validesa

Lectures:

Prieto, G., y Delgado, A.R. (2010). Fiabilidad y validez. *Papeles del Psicólogo*, 31(1), 67-74.

Sireci, S., & Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema*, 26(1), 100-107. (pàgines 100-103)

Test Vàlid - Les inferències que fem a partir del test s'adeqüen al propòsit del test.

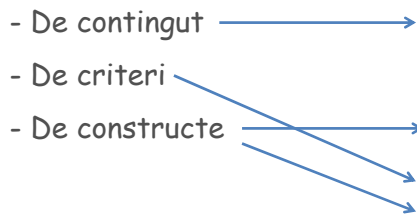
Validesa - Concepte unitari que inclou diverses facetes a validar:

Històric:

- De contingut
- De criteri
- De constructe

Estàndards 1999 i 2014:

- De contingut
- Procés de resposta
- Estructura interna
- Relacions amb altres variables
- Conseqüencial



Facetes a validar:

- Validesa de contingut - domini conductual
- Validesa de criteri - diagnosticar o predir una variable criteri
- Validesa de constructe - concepte psicològic

Validesa de contingut

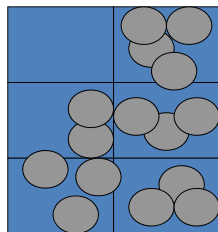
Rellevància

Contingut inclou informació directament relacionada amb el constructe

Representativitat

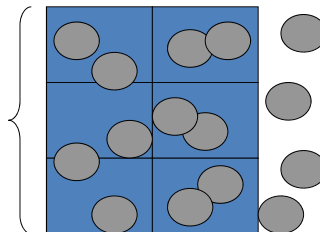
Informació complementària a l'aportada per la resta d'ítems

Falta de representativitat



Definició domini

Falta de rellevància



Figures adaptades de Viladrich et al. (2005). *Psicometria*. Barcelona: Editorial UOC

Validesa de contingut

Passos:

1. Definir i especificar el camp o domini del test.
2. Selecció jutges experts.
3. Procés d'aparellament d'ítems amb els temes d'interès de mesura del test.
4. Valoració de la rellevància i representativitat.
5. Selecció dels ítems adequats.

Exemple: Validesa de contingut

Prova per avaluar els coneixements de Psicometria

Continguts a mostrejar:

TCT / Anàlisi d'ítems / Fiabilitat / Validesa / Baremació / Manual del test

Experts: professors de Psicometria de la UAB

Aparellament: fer correspondre cada ítem amb cada un dels continguts
(p.ex.: ítem 1 -> Baremació, ítem 2 -> Validesa,...)

Valoració: els experts valoren la rellevància i representativitat dels ítems
(p.ex.: puntuant-los de l'1 al 5)

Selecció d'ítems: seleccionar els ítems adequats (més ben valorats pels jutges experts) tenint en compte la importància relativa de cada contingut (p.ex.: 1 ítem de TCT, 5 d'Anàlisi d'ítems,...)

Validesa de criteri

Grau d'eficàcia amb què podem diagnosticar o predir la variable *criteri* a partir de les puntuacions del test.

Tipus:

1. Coeficient de validesa (r_{xy})

Les puntuacions del test i del criteri són quantitatives - correlació entre el test i el criteri

2. Validesa de decisió

Les puntuacions del test i del criteri són dicotòmiques o dicotomitzades - es compara la capacitat de classificació del test respecte a un diagnòstic

Validesa de criteri

Exemples de validesa de criteri

Coeficient de validesa (r_{xy})

Validem...	Criteri que volem predir...	Possible indicador del criteri...
Test d'intel·ligència	Rendiment acadèmic	Notes escolars
Sistema de selecció de personal en una empresa de ventes	Rendiment comercial	Número de ventes setmanals
Prova d'accés a la universitat	Rendiment acadèmic	Nota mitjana de la carrera
Test d'aptitud mecànica	Destresa mecànica	Temps (en hores) que es tarda en arreglar un cotxe amb la mateixa avaria

Dificultat de selecció dels indicadors del criteri

Validesa de criteri

Com decidim quin és l'indicador que cal escollir?

Requisits segons Thorndike i Hagen (1989):

Rellevants. Un indicador és rellevant en la mesura en què es correspon amb el criteri. Cal utilitzar el judici dels jutges experts.

Lliures de biaix. Les mesures del criteri han de representar la verdadera competència dels subjectes i no estar determinades per factors que actuïn diferencialment entre grups. Ex.: secretàries - físic

Fiables. Les mesures del criteri han de ser estables en el temps.

Accessibles. A igualtat de condicions d'adequació, escollirem aquell criteri que és més fàcil i pràctic aconseguir. Tindrem en compte aspectes econòmics, temporals, etc.

Validesa de criteri

Fases:

1. Definir el criteri que es pretén mesurar.
2. Identificar l'indicador o indicadors que s'utilitzaran per mesurar el criteri.
3. Selecció d'una mostra representativa de subjectes.
4. Administració del test i obtenció d'una puntuació per a cada persona.
5. Mesura de cada persona en el criteri: abans del test (retrospectiva), mateix temps que el test (concurrent) o després del test (predictiva).
6. Càlcul coeficient de validesa: determinar el grau de relació entre les puntuacions en el test i en la mesura del criteri.

Criteris valoració validesa concurrent

Prieto, G. y Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 77, 65-75.

- Inadeuada: $|r| < 0.25$
- Suficient: $0.25 \leq |r| < 0.40$
- Bona: $0.40 \leq |r| < 0.50$
- Molt bona: $0.50 \leq |r| < 0.60$
- Excel·lent: $|r| \geq 0.60$

Criteris valoració validesa predictiva

Prieto, G. y Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 77, 65-75.

- Inadeuada: $|r| < 0.20$
- Suficient: $0.20 \leq |r| < 0.35$
- Bona: $0.35 \leq |r| < 0.45$
- Molt bona: $0.45 \leq |r| < 0.55$
- Excel·lent: $|r| \geq 0.55$

Predicció del criteri a partir del test

$$y_i = r_{xy} \frac{S_y}{S_x} (x_i - \bar{x}) + \bar{y}$$

On:

y_i → Puntuació del subjecte en el criteri, es prediu a partir de la puntuació del subjecte en el test

r_{xy} → Coeficient de validesa del test

S_y → Desviació típica del criteri

S_x → Desviació típica del test

x_i → Puntuació del subjecte en el test

\bar{x} → Mitjana del test

\bar{y} → Mitjana del criteri

Interval de confiança del criteri

$$Y_{\text{int}} = y_i \pm z \cdot S_{xy}$$

← Error màxim →

On:

y_i → Puntuació predita del subjecte i

z → Valor de les taules llei normal (95%: $z=1,96$; 99%: $z=2,58$)

S_{xy} → Error típic d'estimació: $S_{xy} = S_y \sqrt{1 - r_{xy}^2}$

On:

S_y → Desviació típica criteri

r_{xy} → Coeficient de validesa del test

Exemple: Validesa de criteri

Suposem que tenim un test amb una fiabilitat de 0,75, una mitjana de 80 i una desviació típica de 16. Es correlaciona amb un criteri de mitjana 20 i una desviació típica de 8, obtenint un coeficient de correlació de 0,46. Un subjecte obté una puntuació de 86 en el test.

Quin és l'error típic d'estimació?
¿En quin interval se situa la puntuació del subjecte en el criteri a un nivell de confiança del 95%?

Exemple: Validesa de criteri

Error típic d'estimació:

$$S_{xy} = 8 \cdot \sqrt{1 - 0,46^2} = 7,10$$

Interval de confiança:

$$y_i = 0,46 \frac{8}{16} (86 - 80) + 20 = 21,38$$

$$Y_{\text{int}} \Rightarrow 21,38 \pm 1,96 \cdot 7,10 \Rightarrow 7,46 \leq Y_{\text{int}} \leq 35,30$$

Exercici: Validesa de criteri

Imaginem que tenim un test amb una fiabilitat de 0,80, una mitjana de 40 i una variància de 64. Es correlaciona amb un criteri de mitjana 60 i una desviació típica de 12, obtenint un coeficient de correlació de 0,52. Un subjecte presenta una puntuació de 35 en el test.

Quin és l'error típic d'estimació?
¿En quin interval se situa la puntuació del subjecte en el criteri a un nivell de confiança del 95%?

Factors que afecten al coeficient de validesa

- Validesa dels ítems
- Variabilitat
- Fiabilitat

$$\text{Si } r_{v_x v_y} = 1 \quad \rightarrow \quad r_{xy} = \sqrt{r_{xx}}$$

Per tant, sempre es compleix que $r_{xy} \leq \sqrt{r_{xx}}$

- Longitud del test

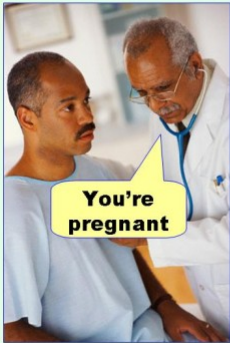

Validesa de decisió

		Diagnòstic criteri		Total
		Positiu	Negatiu	
Resultat test	Positiu	Vertader positiu (f_{11})	Fals positiu (f_{12})	$f_{.1}$
	Negatiu	Fals negatiu (f_{21})	Vertader negatiu (f_{22})	$f_{.2}$
Total		$f_{.1}$	$f_{.2}$	N

Positiu – Amb trastorn – Malalt – Apte – Tractament – Rehabilitació – Teràpia – ...

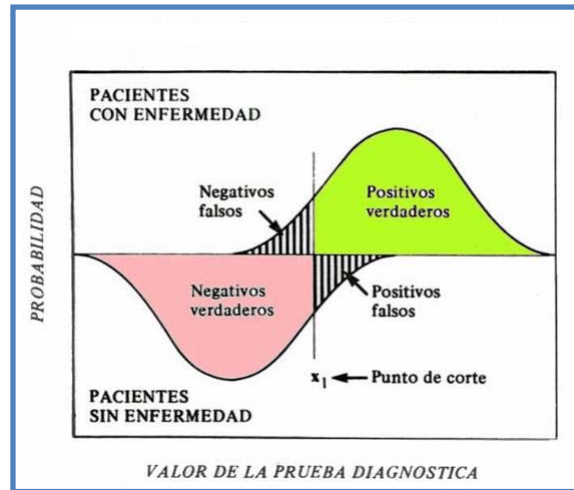
Negatiu – Sense trastorn – Sa – No apte – No tractament – No rehabilitació – No teràpia – ...

Validesa de decisió

Type I error (false positive)	Type II error (false negative)
 <p>A doctor in a white coat is talking to a man in a light blue hospital gown. A yellow speech bubble from the doctor says "You're pregnant".</p>	 <p>A doctor in a white coat is talking to a pregnant woman in a light blue hospital gown. A yellow speech bubble from the doctor says "You're not pregnant".</p>

Imatge extreta de www.marginalrevolution.com

Validesa de decisió



Índexs de validesa de decisió: Proporció de classificacions correctes, sensibilitat i especificitat

$$P_c = \frac{\text{Vertaders positius } (f_{11}) + \text{Vertaders negatius } (f_{22})}{N}$$

$$\text{Sensibilitat} = \frac{\text{Vertaders positius } (f_{11})}{\text{Total persones amb trastorn } (f_{.1})}$$

Capacitat d'un test per a diagnosticar els casos positius. Indica la proporció de persones que tenen el trastorn que han estat classificades correctament. És màxima quan no hi ha FN.

$$\text{Especificitat} = \frac{\text{Vertaders negatius } (f_{22})}{\text{Total persones sanes } (f_{.2})}$$

Capacitat d'un test per a diagnosticar els casos negatius. Indica la proporció de persones que NO tenen el trastorn que han estat classificades correctament. És màxima quan no hi ha FP.

Índexs de validesa de decisió: Coeficient Kappa

$$K = \frac{F_c - F_a}{N - F_a}$$

F_c és la freqüència d'acords

F_a és la freqüència d'acords
per atzar

On:

$$F_c = \text{Vertaders positius } (f_{11}) + \text{Vertaders negatius } (f_{22})$$

$$F_a = \frac{f_{1.} \cdot f_{.1} + f_{2.} \cdot f_{.2}}{N}$$

$$IC \Rightarrow K \pm z \cdot S_{e(K)} \quad S_{e(K)} = \sqrt{\frac{F_a}{N \cdot (N - F_a)}}$$

Valoració dels índexs de validesa de decisió

Valoració de P_c , S y E (Domènech i Granero, 2003)

Valors $\geq 0,80$: Acceptables

Valoració del coeficient Kappa (Altman, 1991)

Coeficient Kappa	Força de la concordança
< 0,20	Pobra
0,21 – 0,40	Dèbil
0,41 – 0,60	Moderada
0,61 – 0,80	Bona
0,81 – 1,00	Molt bona

Domènech, J. M., y Granero, R. (2003). *Anàlisi de dades per a la recerca en Psicologia*. Barcelona: Signo.
Altman, D. G. (1991). *Practical statistics for medical research*. New York: Chapman and Hall.

Exemple: Validesa de decisió

		Depressió Post-part		Total
		Sí	No	
Escala de Hamilton	Sí	125	50	175
	No	25	300	325
Total		150	350	500

$$P_c = \frac{125 + 300}{500} = 0,85$$

$$S = \frac{125}{150} = 0,83$$

$$E = \frac{300}{350} = 0,86$$

$$F_a = \frac{175 \cdot 150 + 325 \cdot 350}{500} = 280$$

$$F_c = 125 + 300 = 425$$

$$K = \frac{425 - 280}{500 - 280} = 0,66$$

$$S_{e(K)} = \sqrt{\frac{280}{500 \cdot (500 - 280)}} = 0,05$$

$$IC(95\%) \Rightarrow 0,66 \pm 1,96 \cdot 0,05$$

$$0,56 \leq K \leq 0,76$$

Exercici: Validesa de decisió

		Diagnòstic clínic	
		No estrès	Estrès
Qüestionari d'estrès	Estrès	200	1235
	No estrès	2400	165

Obtenir:

- Proporció de classificacions correctes
- Sensibilitat
- Especificitat
- Coeficient Kappa i IC (95%; $z = 1,96$)

IMPORTANT!
L'ordre de les columnes està invertit

Exemple: Validesa de decisió - Punt de tall

Disposem de la decisió emesa per un psicòleg de la necessitat de fer teràpia (TE) o no fer-ne (NT) dirigides a 12 persones.

Se'ls aplica una escala les puntuacions de la qual oscil·len entre 1 i 10.

Com podem determinar aquell punt de tall que optimitzi les seves classificacions fetes per l'escala?

Persones	Escala	Diagnòstic
A	6	NT
B	6	TE
C	7	TE
D	8	NT
E	5	NT
F	8	TE
G	4	NT
H	9	TE
I	3	NT
J	7	TE
K	7	NT
L	10	TE

Exemple: Validesa de decisió - Punt de tall 1

Persones	Escala	Diagnòstic	≥ 1	
A	6	NT	TE	FP
B	6	TE	TE	V _p
C	7	TE	TE	V _p
D	8	NT	TE	FP
E	5	NT	TE	FP
F	8	TE	TE	V _p
G	4	NT	TE	FP
H	9	TE	TE	V _p
I	3	NT	TE	FP
J	7	TE	TE	V _p
K	7	NT	TE	FP
L	10	TE	TE	V _p

Exemple: Validesa de decisió - Punt de tall 2

Persones	Escala	Diagnòstic	≥ 2	
A	6	NT	TE	FP
B	6	TE	TE	V _p
C	7	TE	TE	V _p
D	8	NT	TE	FP
E	5	NT	TE	FP
F	8	TE	TE	V _p
G	4	NT	TE	FP
H	9	TE	TE	V _p
I	3	NT	TE	FP
J	7	TE	TE	V _p
K	7	NT	TE	FP
L	10	TE	TE	V _p

Exemple: Validesa de decisió - Punt de tall 3

Persones	Escala	Diagnòstic	≥ 3	
A	6	NT	TE	FP
B	6	TE	TE	V _p
C	7	TE	TE	V _p
D	8	NT	TE	FP
E	5	NT	TE	FP
F	8	TE	TE	V _p
G	4	NT	TE	FP
H	9	TE	TE	V _p
I	3	NT	TE	FP
J	7	TE	TE	V _p
K	7	NT	TE	FP
L	10	TE	TE	V _p

Exemple: Validesa de decisió - Punt de tall 4

Persones	Escala	Diagnòstic	≥ 4	
A	6	NT	TE	FP
B	6	TE	TE	V _p
C	7	TE	TE	V _p
D	8	NT	TE	FP
E	5	NT	TE	FP
F	8	TE	TE	V _p
G	4	NT	TE	FP
H	9	TE	TE	V _p
I	3	NT	NT	V _N
J	7	TE	TE	V _p
K	7	NT	TE	FP
L	10	TE	TE	V _p

Exemple: Validesa de decisió - Punt de tall 5

Persones	Escala	Diagnòstic	≥ 5	
A	6	NT	TE	FP
B	6	TE	TE	V _p
C	7	TE	TE	V _p
D	8	NT	TE	FP
E	5	NT	TE	FP
F	8	TE	TE	V _p
G	4	NT	NT	V _N
H	9	TE	TE	V _p
I	3	NT	NT	V _N
J	7	TE	TE	V _p
K	7	NT	TE	FP
L	10	TE	TE	V _p

Exemple: Validesa de decisió - Punt de tall 6

Persones	Escala	Diagnòstic	≥ 6	
A	6	NT	TE	FP
B	6	TE	TE	V _p
C	7	TE	TE	V _p
D	8	NT	TE	FP
E	5	NT	NT	V _N
F	8	TE	TE	V _p
G	4	NT	NT	V _N
H	9	TE	TE	V _p
I	3	NT	NT	V _N
J	7	TE	TE	V _p
K	7	NT	TE	FP
L	10	TE	TE	V _p

Exemple: Validesa de decisió - Punt de tall 7

Persones	Escala	Diagnòstic	≥ 7	
A	6	NT	NT	V _N
B	6	TE	NT	FN
C	7	TE	TE	V _p
D	8	NT	TE	FP
E	5	NT	NT	V _N
F	8	TE	TE	V _p
G	4	NT	NT	V _N
H	9	TE	TE	V _p
I	3	NT	NT	V _N
J	7	TE	TE	V _p
K	7	NT	TE	FP
L	10	TE	TE	V _p

Exemple: Validesa de decisió - Punt de tall 8

Persones	Escala	Diagnòstic	≥ 8	
A	6	NT	NT	V_N
B	6	TE	NT	FN
C	7	TE	NT	FN
D	8	NT	TE	FP
E	5	NT	NT	V_N
F	8	TE	TE	V_p
G	4	NT	NT	V_N
H	9	TE	TE	V_p
I	3	NT	NT	V_N
J	7	TE	NT	FN
K	7	NT	NT	V_N
L	10	TE	TE	V_p

Exemple: Validesa de decisió - Punt de tall 9

Persones	Escala	Diagnòstic	≥ 9	
A	6	NT	NT	V_N
B	6	TE	NT	FN
C	7	TE	NT	FN
D	8	NT	NT	V_N
E	5	NT	NT	V_N
F	8	TE	NT	FN
G	4	NT	NT	V_N
H	9	TE	TE	V_p
I	3	NT	NT	V_N
J	7	TE	NT	FN
K	7	NT	NT	V_N
L	10	TE	TE	V_p

Exemple: Validesa de decisió - Punt de tall 10

Persones	Escala	Diagnòstic	≥ 10	
A	6	NT	NT	V_N
B	6	TE	NT	FN
C	7	TE	NT	FN
D	8	NT	NT	V_N
E	5	NT	NT	V_N
F	8	TE	NT	FN
G	4	NT	NT	V_N
H	9	TE	NT	FN
I	3	NT	NT	V_N
J	7	TE	NT	FN
K	7	NT	NT	V_N
L	10	TE	TE	V_P

Exemple: Validesa de decisió - Punt de tall

Per cada possible punt de tall, calcular el número de FP, FN i el número total d'errors

Escollir el punt de tall que minimitzi el número d'errors

En cas d'empat, valorar la implicació de cometre un FP o un FN



Punt de Tall	Fals Positiu	Fals Negatiu	Total errors
1	6	0	6
2	6	0	6
3	6	0	6
4	5	0	5
5	4	0	4
≥ 6	3	0	3
7	2	1	3
8	1	3	4
9	0	4	4
10	0	5	5

Exemple: Validesa de decisió - Punt de tall

Agafant com a punt de tall ≥ 6 , obtindríem...

		Psicòleg		Total
		TE	NT	
Escala	TE	6	3	9
	NT	0	3	3
Total		6	6	12

I podríem calcular els índexs de validesa de decisió per aquest punt de tall concret

Exercici: Validesa de decisió - Punt de tall

En una residència d'avis es posa a prova la validesa d'una escala d'observació dissenyada per a detectar la dependència funcional dels residents i assignar-los a un grup de rehabilitació (R) o no (NR).

Es disposa de les puntuacions obtingudes per 11 residents en l'escala d'observació, així com del diagnòstic emès per un especialista (R o NR).

Determinar el punt de tall òptim en l'escala d'observació, tenint en compte que la rehabilitació no es considera nociva per l'individu. Construir la corresponent taula de classificacions.

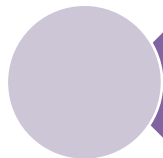
Exercici: Validesa de decisió - Punt de tall

Residents	Escala observació	Diagnòstic especialista
1	7	NR
2	2	NR
3	1	NR
4	1	NR
5	5	NR
6	6	R
7	4	R
8	4	NR
9	3	NR
10	9	R
11	10	R

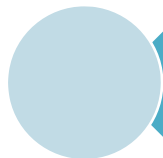
Validesa de constructe

Tracta d'inferir el grau en què el test mesura el concepte psicològic (constructe) que pretén mesurar.

Es basa en evidències múltiples i contingents:



Anàlisi factorial - identificació (AFE) o confirmació (AFC) dels trets psicològics que mesura el test



Matriu multitret-multimètode (Campbell i Fiske, 1959)
Validesa convergent (mateix tret - diferent mètode)
Validesa discriminant (diferent tret - mateix mètode)

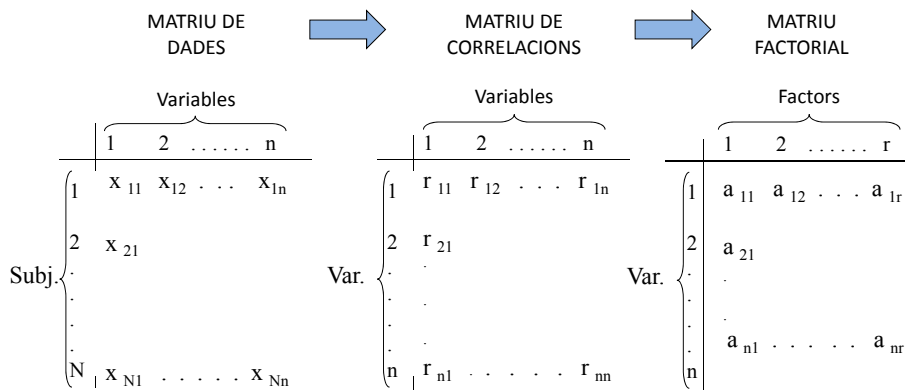
Validesa de constructe

Passos:

1. Definir i formular hipòtesis sobre el constructe.
2. Recollir dades per comprovar les hipòtesis.
3. Consistència de les dades amb les hipòtesis formulades o altres alternatives.

Validesa de constructe: Anàlisi factorial

Tècnica de reducció de dades:
 n variables \rightarrow r factors ($r < n$)



Validesa de constructe: Anàlisi factorial

Suposem que construim un test per mesurar tres factors de manera que:

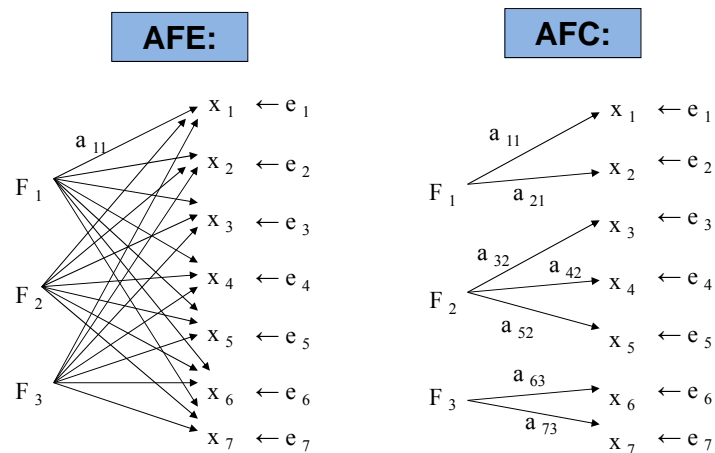
Factor 1: ítems 1 i 2 Factor 2: ítems 3, 4 i 5 Factor 3: ítems 6 i 7

Estructura matriu factorial:

		AFE:					AFC:		
		factors					factors		
		1	2	3			1	2	3
ítems	1	a_{11}	a_{12}	a_{13}	ítems	1	a_{11}	0	0
	2	a_{21}	a_{22}	a_{23}		2	a_{21}	0	0
	3	a_{31}	a_{32}	a_{33}		3	0	a_{32}	0
	4	a_{41}	a_{42}	a_{43}		4	0	a_{42}	0
	5	a_{51}	a_{52}	a_{53}		5	0	a_{52}	0
	6	a_{61}	a_{62}	a_{63}		6	0	0	a_{63}
	7	a_{71}	a_{72}	a_{73}		7	0	0	a_{73}

Validesa de constructe: Anàlisi factorial

Representació gràfica



Exemple: Anàlisi factorial exploratòria

"Se efectuó un análisis factorial exploratorio mediante extracción de ejes principales y rotación oblimin. Teniendo en cuenta el elevado número de ítems que se desarrollaron al inicio, se depuró la escala eliminando todos aquellos ítems que presentaran saturaciones inferiores a 0,35 o saturaciones complejas (superiores a 0,35 en más de un factor)."

Fragment extret de Boada Grau y cols., (2009) CONS-32: estructura factorial y propiedades psicométricas. *Psicothema*, 21(1), 165-169.

Exemple: Anàlisi factorial exploratòria

Matriz de saturaciones de los factores de la escala CONS-32

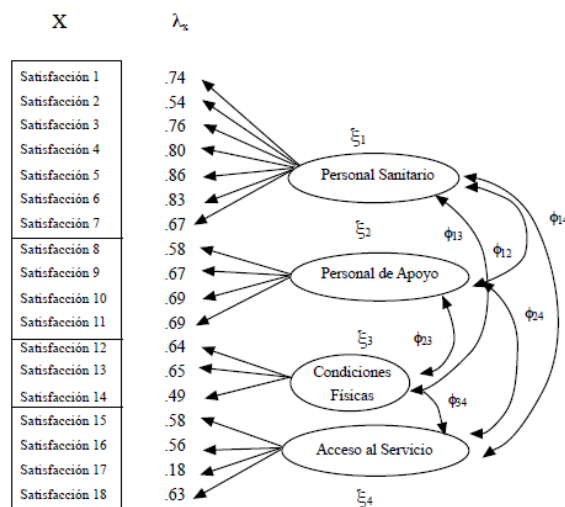
Ítems	F1	F2	F3	F4
2. En el interior de la obra hay escaleras que permiten moverte de un modo seguro	0.60	0.00	0.20	0.06
3. Conozco las normas de seguridad de la obra	0.63	0.08	0.04	-0.07
5. Los andamios y las maderas están en mal estado	0.55	0.13	0.23	0.06
15. Se colocan los dispositivos para evitar la caída de personas	0.63	-0.17	0.03	0.08
16. Las zanjas están entibadas y protegidas	0.51	-0.17	0.08	-0.10
17. Los andamios tienen barandillas	0.71	-0.05	-0.21	0.08
22. Las escaleras tienen barandillas	0.64	0.02	0.01	0.13
30. Tenemos las medidas de seguridad a nuestro alcance (por ejemplo, redes para protegerse de la caída, arneses para sujetarnos, cascos, guantes, etc.)	0.58	-0.09	-0.02	0.08
4. Hago soldaduras con traje protector	0.19	-0.46	0.02	0.03
6. Al cortar madera uso guantes	-0.07	-0.74	0.06	-0.02
14. Al llegar a la obra por la mañana se revisa que la grúa, los andamios y el resto del material siga en perfecto estado	0.24	-0.42	0.09	0.15
8. Al cortar madera uso mascarilla de respiración	-0.07	-0.72	-0.08	0.02
25. Cuando trabajo con el martillo mecánico para taladrar el suelo, me protejo los oídos con cascos	0.21	-0.51	0.07	0.05
28. Uso gafas protectoras para cortar vidrio	-0.02	-0.58	0.20	0.03
31. Cuando corto baldosas utilizo guantes	0.02	-0.60	-0.02	-0.12
32. Entro en la obra con el casco puesto	0.33	-0.53	-0.07	0.02

Exemple: Anàlisi factorial confirmatòria

“El primero de los objetivos era comprobar si la estructura factorial de la escala era la esperada, con las 4 dimensiones descritas anteriormente. Para ello se llevó a cabo un análisis factorial confirmatorio mediante el programa AMOS 4. Los parámetros del modelo original fueron estimados siguiendo el criterio de Máxima Verosimilitud.”

Fragment extret de Varela Mallou, Rial Boubeta y García Cueto (2003). Presentación de una Escala de Satisfacción con los Servicios Sanitarios de Atención Primaria. *Psicothema*, 15(4), 656-661.

Exemple: Anàlisi factorial confirmatòria



Exemple: Anàlisi factorial confirmatòria

Ítems	Factor I (Extraversió)	Factor II (Neuroticisme)	Factor III (Paranoidisme)
Extraversió 1	0,80	0,00	0,00
Extraversió 2	0,90	0,00	0,00
Extraversió 3	0,70	0,00	0,00
Extraversió 4	0,80	0,00	0,00
Neuroticisme 1	0,00	0,75	0,00
Neuroticisme 2	0,00	0,82	0,00
Neuroticisme 3	0,00	0,90	0,00
Paranoidisme 1	0,00	0,00	0,85
Paranoidisme 2	0,00	0,00	0,92
Paranoidisme 3	0,00	0,00	0,70

Exercici: AFE i AFC

Escala de Insomnio de Atenas

Edad:.....

Sexo:.....

Fecha:.....

Instrucciones: Esta escala pretende registrar su propia valoración acerca de cualquier dificultad que haya podido experimentar. Por favor, responda los siguientes ítems (marcando con un círculo el número apropiado), indicando las dificultades que le hayan ocurrido como mínimo tres veces a la semana durante el último mes.

1. Inducción del sueño (tiempo que tarda en dormirse después de apagar la luz)

0: Ningún problema 1: Ligeramente retrasado 2: Marcadamente retrasado 3: Muy retrasado o no durmió

2. Despertares durante la noche

0: Ningún problema 1: Problema menor 2: Problema considerable 3: Problema serio o no durmió

3. Despertar final antes de lo deseado

0: No fue antes 1: Un poco antes 2: Notablemente antes 3: Mucho antes o no durmió

4. Duración total del sueño

0: Suficiente 1: Ligeramente insuficiente 2: Notablemente insuficiente 3: Muy insuficiente o no durmió

5. Calidad general del sueño (no importa cuánto tiempo durmió)

0: Satisfactoria 1: Ligeramente insatisfactoria 2: Notablemente insatisfactoria 3: Muy insatisfactoria o no durmió

6. Sensación de bienestar durante el día

0: Normal 1: Ligeramente disminuida 2: Notablemente disminuida 3: Muy disminuida

7. Funcionamiento (físico y mental) durante el día

0: Normal 1: Ligeramente disminuido 2: Notablemente disminuido 3: Muy disminuido

8. Somnolencia durante el día

0: Ninguna 1: Moderada 2: Considerable 3: Intensa

El periodo de la auto-administración puede variar, dependiendo del diseño de cada estudio. Cuando la auto-administración se haya efectuado en un periodo distinto al del mes anterior, la segunda frase de las instrucciones debería corregirse consecuentemente.

Exercici: AFE i AFC

Tenint en compte que...

L'Escala d'Insomni d'Atenes pretén mesurar les dificultats en el dormir en dos moments diferents: a) aquells problemes que sorgeixen durant la nit, i b) aquells problemes que sorgeixen durant el dia

I analitzant l'enunciat dels ítems del test...

Proposar una hipotètica matriu factorial resultant d'una AFE i una altra resultant d'una AFC

Anàlisi factorial exploratòria (AFE)

Ítem	Sexe	AIS_1	AIS_2	AIS_3	AIS_4	AIS_5	AIS_6	AIS_7	AIS_8
1	19	Home	0	0	1	0	1	0	1
2	26	Home	0	1	1	1	1	1	0
3	19	Dona	2	0	0	2	2	2	1
4	20	Dona	0	1	0	1	0	1	0
5	Home	2	1	2	1	2	1	2	1
6	20	Dona	0	0	0	0	0	0	0
7	19	Dona	0	0	1	2	2	3	2
8	35	Dona	0	1	2	2	2	1	1
9	22	Dona	0	1	0	0	0	0	0
10	20	Dona	1	0	0	0	0	0	1
11	19	Dona	0	0	1	0	0	0	1
12	23	Dona	1	1	0	1	1	1	2
13	19	Home	0	1	0	1	1	0	1
14	19	Dona	0	1	1	1	1	0	1
15	22	Dona	0	0	1	0	0	0	0
16	20	Home	2	1	0	2	2	1	0
17	19	Dona	0	0	0	1	0	0	0
18	20	Home	0	1	0	1	1	0	0
19	18	Dona	0	0	0	0	0	0	0
20	19	Dona	0	0	1	0	0	0	1
21	19	Dona	0	1	1	1	0	0	0
22	24	Home	0	0	1	0	0	0	0
23	19	Dona	0	0	0	0	0	0	0
24	19	Home	0	0	0	1	0	0	1
25	20	Home	2	1	0	1	2	1	2
26	21	Dona	0	1	0	1	0	0	1
27	21	Home	1	1	0	1	0	1	1
28	19	Dona	1	1	0	0	1	1	0
29	25	Home	0	0	0	0	0	0	0
30	20	Dona	1	1	0	0	0	1	1
31	19	Home	2	1	0	2	1	1	2
32	39	Dona	0	1	1	1	1	1	2
33	25	Dona	0	1	0	1	0	1	1
34	20	Dona	0	0	0	0	1	1	2
35	19	Dona	0	1	1	2	1	2	2
36	21	Dona	2	1	1	2	1	1	1
37	20	Home	0	0	0	0	0	0	0
38	19	Dona	1	1	1	1	0	1	2
39	19	Dona	0	0	0	1	0	1	1
40	19	Home	0	1	0	1	1	0	0
41	18	Home	1	1	0	1	1	1	0
42	19	Home	1	1	0	1	1	1	1

Matriu de dades

		AIS_1	AIS_2	AIS_3	AIS_4	AIS_5	AIS_6	AIS_7	AIS_8
AIS_1	Correlació de Pearson	1	.394*	.207	.337	.405*	.347*	.239	.273*
	Sig. (bilateral)		.000	.015	.000	.000	.000	.005	.001
	N	137	137	137	137	137	137	136	137
AIS_2	Correlació de Pearson	.394*	1	.208*	.165	.216*	.144	.052	.061
	Sig. (bilateral)	.000		.015	.055	.012	.063	.551	.290
	N	137	137	137	137	137	137	136	137
AIS_3	Correlació de Pearson	.207	.208*	1	.076	.299*	.163	.140	.001
	Sig. (bilateral)	.015	.015		.378	.000	.058	.104	.967
	N	137	137	137	137	137	137	136	137
AIS_4	Correlació de Pearson	.337*	.165	.076	1	.447*	.484*	.422*	.501*
	Sig. (bilateral)	.000	.055	.378		.000	.000	.000	.000
	N	137	137	137	137	137	137	136	137
AIS_5	Correlació de Pearson	.405*	.216*	.299*	.447*	1	.642*	.483*	.474**
	Sig. (bilateral)	.000	.012	.000	.000		.000	.000	.000
	N	137	137	137	137	137	137	136	137
AIS_6	Correlació de Pearson	.347*	.144	.163	.484*	.642*	1	.761*	.603*
	Sig. (bilateral)	.000	.093	.058	.000	.000		.000	.000
	N	137	137	137	137	137	137	136	137
AIS_7	Correlació de Pearson	.239	.052	.140	.422*	.483*	.483*	1	.492**
	Sig. (bilateral)	.005	.551	.104	.000	.000	.000		.000
	N	136	136	136	136	136	136	136	136
AIS_8	Correlació de Pearson	.273*	.061	.001	.501*	.474**	.603*	.492**	1
	Sig. (bilateral)	.001	.290	.987	.000	.000	.000	.000	
	N	137	137	137	137	137	137	136	137

** La correlació és significativa al nivell 0.01 (bilateral).

* La correlació és significativa al nivell 0.05 (bilateral).

Matriu de correlacions

Matriu de estructura

	Componente	
	1	2
AIS_1	.425	.704
AIS_2	.121	.774
AIS_3	.129	.636
AIS_4	.703	.260
AIS_5	.742	.469
AIS_6	.693	.250
AIS_7	.614	.122
AIS_8	.786	.082

Matriu factorial

Anàlisi factorial exploratòria (AFE)

Passos:

1. Comprovar condicions d'aplicació: Kaiser-Meyer-Olkin (KMO) i Prova d'esfericitat de Barlett.
2. Seleccionar un mètode d'extracció de factors.
3. Escollir un mètode de rotació (ortogonal o obliqua).
4. Interpretar els resultats: comunalitats, nombre de factors, variància explicada, matriu factorial, etc.

Exemple AFE amb SPSS (Athens Insomnia Scale)

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		.803
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	372,283
	gl	28
	Sig.	.000

KMO (Kaiser-Mayer-Olkin) indica fins a quin punt les associacions entre variables (ítems) poden ser explicades per un número de factors més reduït que el de les variables.

Prova d'esfericitat de Bartlett: avalua l' H_0 de què les variables o ítems no estan relacionats entre ells.

Comunalidades

	Inicial	Extracción
AIS_1	1,000	,558
AIS_2	1,000	,606
AIS_3	1,000	,406
AIS_4	1,000	,501
AIS_5	1,000	,632
AIS_6	1,000	,797
AIS_7	1,000	,671
AIS_8	1,000	,634

Comunalitats: quantitat d'informació que l'ítem comparteix amb els factors comuns.

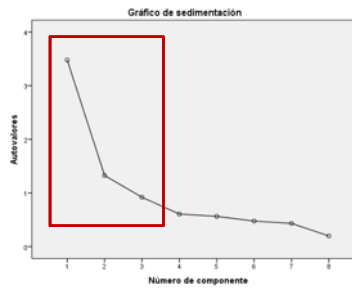
Exemple AFE amb SPSS (Athens Insomnia Scale)

Críteri de Kaiser
 (valor propi > 1)

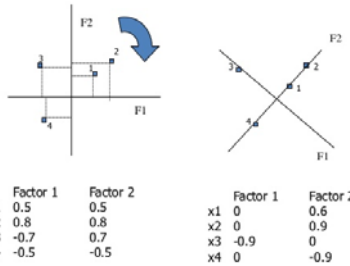
Variància total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación
	Total	% de la variancia	% acumulado	Total	% de la variancia	% acumulado	
1	3,479	43,489	43,489	3,479	43,489	43,489	3,333
2	1,325	16,569	60,058	1,325	16,569	60,058	1,870
3	,920	11,495	71,552				
4	,607	7,591	79,143				
5	,563	7,039	86,182				
6	,476	5,947	92,129				
7	,432	5,397	97,526				
8	,199	2,474	100,000				

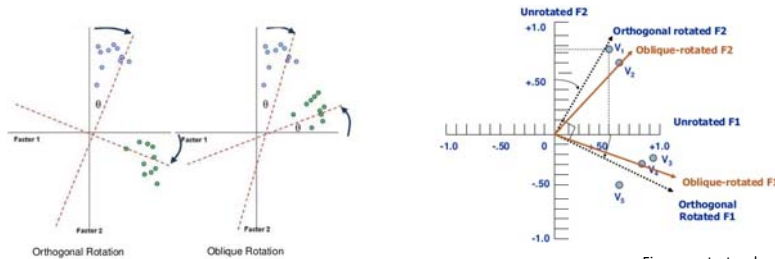
Gràfic de sedimentació



AFE: Rotació



Tipus de rotació: rotació ortogonal vs obliqua



Figures extretes de slideshare.net

Exemple AFE amb SPSS (Athens Insomnia Scale)

Matriu inicial

Matriz de componentes^a

	Componente	
	1	2
AIS_1	,570	,483
AIS_2	,317	,711
AIS_3	,287	,569
AIS_4	,700	-,106
AIS_5	,791	,083
AIS_6	,867	-,215
AIS_7	,761	-,302
AIS_8	,726	-,328

Matriu rotada (rotació obliqua)

Matriz de estructura

	Componente	
	1	2
AIS_1	,425	,704
AIS_2	,121	,774
AIS_3	,129	,636
AIS_4	,703	,260
AIS_5	,742	,469
AIS_6	,893	,250
AIS_7	,814	,122
AIS_8	,786	,082

Matriz de correlaciones de componentes

Componente	1	2
1	1,000	,429
2	,429	1,000

Comunalitat AIS_1

$$0,570^2 + 0,483^2 = 0,558$$

Autovalor Factor 1
(abans de la rotació)

$$0,570^2 + 0,317^2 + \dots + 0,761^2 + 0,726^2 = 3,481$$

Autovalor Factor 1
(després de la rotació)

$$0,425^2 + 0,121^2 + \dots + 0,814^2 + 0,786^2 = 3,335$$

Resum AFE

Condicions d'aplicació:

- Kaiser-Meyer-Olkin (KMO)
- Prova d'esfericitat de Barlett

Mètode d'extracció:

- Eixos principals
- Mínims quadrats
- Màxima versemblança
- Etc.

Rotació:

- Ortogonal (Varimax,...)
- Obliqua (Oblimin,...)

Determinar el nombre de factors:

- Criteri de Kaiser (valor propi >1)
- Gràfic de sedimentació



Interpretar resultats

Resum AFE

Correlació ítem - factor

Ítems	Factor 1	
	Saturació/càrrega factorial	Comunalitat
Ítem 1	0.767	0.588
Ítem 2	0.790	0.624
Ítem 3	0.504	0.254
Ítem 4	0.304	0.092
Autovalor		1.559
Percentatge de variància explicada		38.97%

Proporció de variància que el factor explica de l'ítem

El factor explica el 58.8% de la variància de l'ítem 1

El total de variància que el factor explica del conjunt d'ítems (Σ del quadrat de les saturacions)

El factor explica el 38.97% de la variància dels ítems $(100 \cdot 1.559)/4$

Validesa de constructe: Matriu multitret-multimètode

	A1	B1	C1	A2	B2	C2	A3	B3	C3
A1	(0.89)								
B1	0.51	(0.88)							
C1	0.38	0.37	(0.76)						
A2	(....)					
B2	(....)				
C2	(....)			
A3	(....)		
B3	(....)	
C3	(....)

A, B, C són trets
 1, 2, 3 són mètodes

(...) Valors monotret-monomètode (FIABILITAT)

... Valors heterotret-monomètode (V. DISCRIMINANT)

... Valors monotret-heteromètode (V. CONVERGENT)

... Valors heterotret-heteromètode

Validesa de constructe: Matriu multitret-multimètode

	A1	B1	C1	A2	B2	C2	A3	B3	C3
A1	0.89								
B1	0.03	0.88							
C1	0.28	0.17	0.76						
A2	0.73	0.14	0.22	0.93					
B2	0.15	0.69	0.03	0.18	0.79				
C2	0.12	0.04	0.81	0.03	0.15	0.85			
A3	0.77	0.11	0.09	0.77	0.21	0.16	0.83		
B3	0.21	0.75	0.18	0.21	0.68	0.03	0.15	0.92	
C3	0.19	0.05	0.78	0.22	0.09	0.72	0.14	0.09	0.89

A, B, C són trets
1, 2, 3 són mètodes

Exemple: Matriu multitret-multimètode

- En la següent matriu multitret-multimètode hi ha tres trets:
 - Extraversió
 - Lideratge
 - Intel·ligència social
- Cada tret està mesurat amb tres mètodes:
 - Autoinforme (AI)
 - Observació sistemàtica (OS)
 - Entrevista personal (EP)

Exemple: Matriu multitret-multimètode

		Extraversió			Lideratge			Intel·ligència social		
		AI	OS	EP	AI	OS	EP	AI	OS	EP
Extraversió	AI	0.80								
	OS	0.70	0.80							
	EP	0.60	0.70	0.90						
Lideratge	AI	0.20	0.21	0.10	0.80					
	OS	0.15	0.09	0.08	0.70	0.80				
	EP	0.10	0.12	0.10	0.60	0.60	0.80			
Intel·ligència social	AI	0.20	0.10	0.25	0.10	0.10	0.30	0.90		
	OS	0.09	0.10	0.10	0.10	0.20	0.10	0.70	0.80	
	EP	0.22	0.13	0.20	0.30	0.10	0.20	0.70	0.60	0.80

Exemple: Matriu multitret-multimètode

- A la diagonal principal apareixen les correlacions dels tests amb sí mateixos, això ens indica, els coeficients de **fiabilitat**.
- En l'anàlisi de la **validesa convergent**, és a dir en el cas que un mateix tret es mesura amb diferents mètodes, s'obtenen valors iguals o superiors a 0.60 en tots els casos.
- Les correlacions entre les mesures de diferents trets mesurades amb el mateix mètode ens informen de la **validesa discriminant**. En l'exemple la màxima correlació és de 0.20.

Exercici: Matriu multitret-multimètode

- Mesurem tres trets:
 - Raonament numèric (RN)
 - Factors espacial (FE)
 - Raonament abstracte (RA)

- Cada tret està mesurat amb tres mètodes:
 - Ítems V/F (VF)
 - Ítems d'elecció múltiple (EM)
 - Ítems de format incomplet (FI)

Exemple extret de Barbero, Vila y Suárez (2006). *Psicometría*. Madrid: UNED.

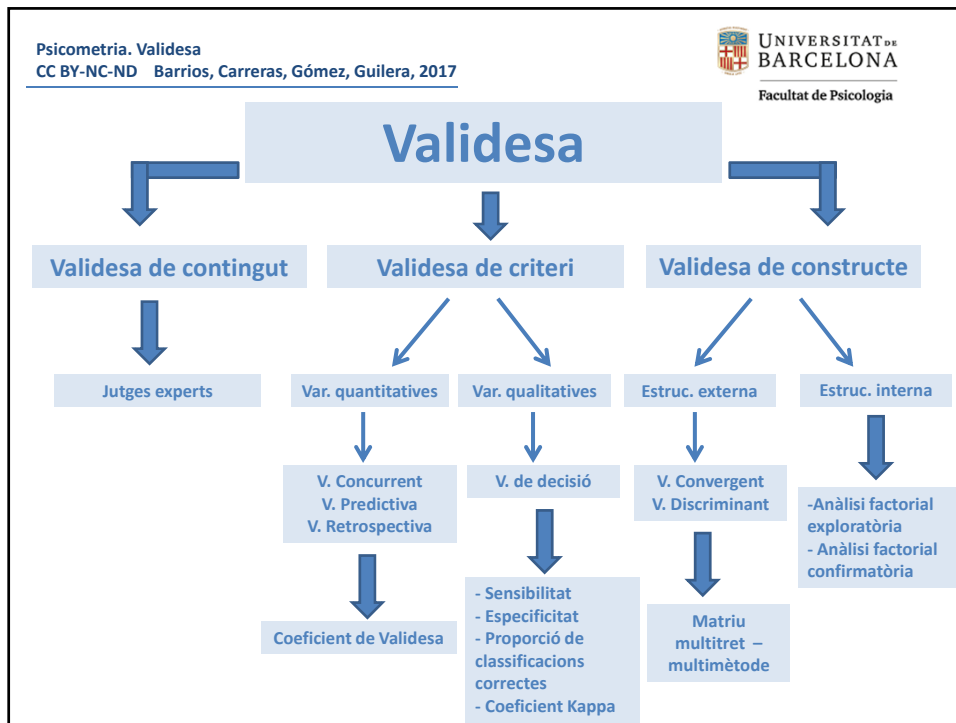
Exercici: Matriu multitret-multimètode

		Veritable /Fals			Elecció múltiple			Format incomplet		
		RN	FE	RA	RN	FE	RA	RN	FE	RA
Veritable/Fals	RN	0.95								
	FE	0.20	0.90							
	RA	0.30	0.28	0.92						
Elecció múltiple	RN	0.90	0.31	0.40	0.93					
	FE	0.26	0.87	0.33	0.37	0.94				
	RA	0.43	0.20	0.84	0.26	0.37	0.88			
Format incomplet	RN	0.79	0.27	0.31	0.77	0.15	0.23	0.89		
	FE	0.11	0.68	0.22	0.24	0.67	0.31	0.19	0.93	
	RA	0.19	0.18	0.50	0.19	0.33	0.72	0.41	0.30	0.64

Exercici: Matriu multitret-multimètode

- Identificar els coeficients de fiabilitat
- Identificar els valors monotret-heteromètode
- Identificar els valors multitret-monomètode

- Valorar la validesa convergent i la validesa discriminant



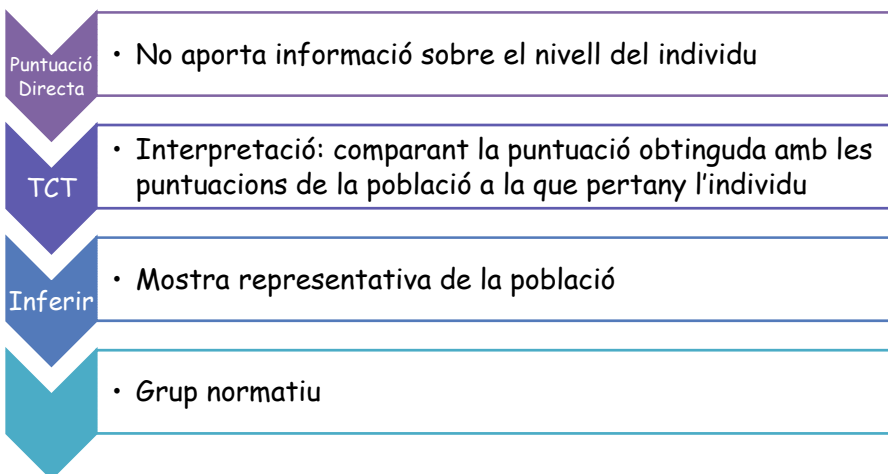
5. Baremació

TEMA: Baremació

Lectura:

Gómez Benito, J. (1997). *Construcció d'instruments de mesura*. Barcelona: Edicions de la Universitat Oberta de Catalunya [únicament de la pàgina 34 a la 38]

Transformació de les Puntuacions



Fases d'un Estudi Normatiu

1. Identificació de la població d'interès.
2. Selecció d'una mostra representativa → Grup normatiu.
3. Recollida de les dades → Aplicació del test.
4. Transformació de puntuacions directes en normatives.
5. Descripció de l'estudi normatiu en el manual del test.

Té caducitat: canvis socioculturals
amb el pas del temps

Cal recalcular els barems per
adaptar el test a grups diferents

Informació que ha de contenir el manual d'un test sobre les normes d'interpretació.

La mostra, el disseny mostral, les taxes de participació i els estadístics descriptius

Permetre, a l'usuari del test, d'avaluar l'adequació del procés i interpretar els resultats obtinguts

Si el test va dirigit a grups diferents, cal definir les normes per a cada grup

Tipus de Transformacions

- Normes Cronològiques.
- Percentils.
- Puntuacions típiques.
- Puntuacions típiques normalitzades.
- Puntuacions típiques derivades.

Normes Cronològiques

Determinar l'EM de cada EC en el test (Binet principis segle).

$$QI = \frac{EM}{EC} \cdot 100$$

- Només es poden utilitzar quan estudiem una característica que experimenta canvis progressius amb l'edat.
- Molt utilitzades en les primeres escales evolutives.

Fases Normes Cronològiques

1. Seleccionar mostres de nens corresponents a les diferents edats a les que el test va dirigit.
2. S'aplica el test als nens de cada edat, i es calcula la seva puntuació mitjana en el test per cada edat.
3. Es construeix una taula en la que a cada edat se li assigna la puntuació mitjana corresponent en el test.

Exemple: Normes Cronològiques

Tenint en compte la taula següent:

Edat cronològica (EC)	Mitjana de les PD	Edat mental (EM)
9	25	9
10	30	10
11	32	11

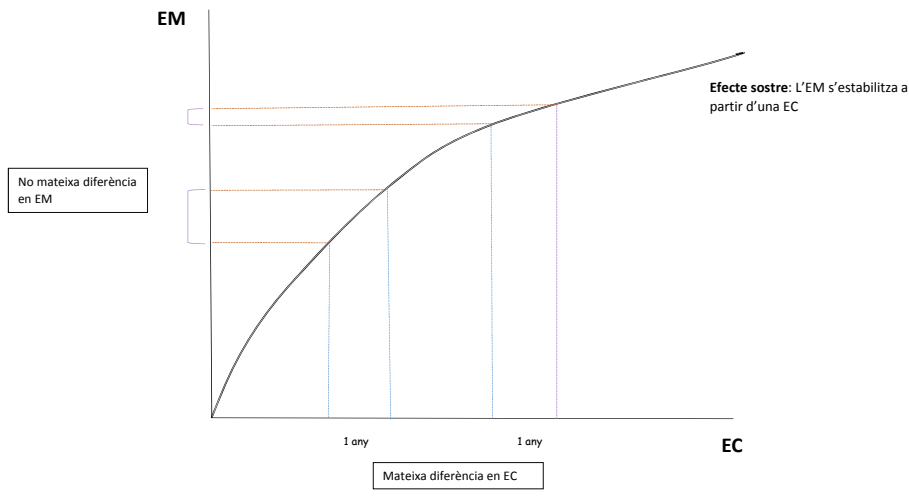
Si apliquem el test a un nen de 12 anys i obté una puntuació de 30, quina EM li correspon? *RESPOSTA:* Una EM de 10

En cas que l'edat d'aquest nen fos de 9, quin QI li correspon? *RESPOSTA:* Un QI de 111
$$QI = \frac{10}{9} \cdot 100 = 111$$

Normes Cronològiques: Inconvenients

- La relació entre EM i EC no és una unitat uniforme \Rightarrow Les diferències obtingudes en EM no tenen el mateix significat per a totes les EC
- Les escales a les que s'apliquen tenen una edat màxima a partir de la qual la puntuació mitjana no augmenta amb l'edat

Normes Cronològiques: Inconvenients



Percentils

Posició d'un individu en la variable que mesura el test respecte al seu grup de referència (grup normatiu).

$$P = \frac{F_{i-1} + 0.5 f_i}{N} \cdot 100$$

➤ On:

- F_{i-1} - la freqüència acumulada per sota de la PD(x).
- f_i - la freqüència de persones que tenen la PD(x).
- N - nombre de persones del grup normatiu.

➤ Avantatge: Facilitat d'interpretació

➤ Inconvenient: Escala ordinal (lloc ↔ diferència)

Percentils

Diferències iguals entre percentils no impliquen diferències iguals entre puntuacions directes

PD	P
1	1
2	5
3	12
4	15
5	19

1 punt de diferència en PD → 7 punts de diferència en P (entre PD 1 i 2)

2 punts de diferència en PD → 7 punts de diferència en P (entre PD 3 i 5)

Exemple: Percentils

x	f_i	F_i	P	P
0	4	4	0.25	1
1	15	19	1.4375	1
2	13	32	3.1875	3
3	24	56	5.5	6
4	63	119	10.9375	11
5	81	200	19.9375	20
6	93	293	30.8125	31
7	85	378	41.9375	42
8	99	477	53.4375	53
9	108	585	66.375	66
10	75	660	77.8125	78
11	31	691	84.4375	84
12	29	720	88.1875	88
13	41	761	92.5625	93
14	27	788	96.8125	97
15	12	800	99.25	99

$$P = \frac{F_{i-1} + 0.5 f_i}{N} \cdot 100$$

$$0 \rightarrow P = \frac{0 + 0.5 \cdot 4}{800} \cdot 100 = 0.25 \rightarrow 1$$

$$1 \rightarrow P = \frac{4 + 0.5 \cdot 15}{800} \cdot 100 = 1.44 \rightarrow 1$$

Exercici: Percentils

Es passa un test de 10 ítems a una mostra de 200 individus, obtenint-se la distribució de freqüències que es detalla:

x	f_i	F_i	P	P
0	4			
1	6			
2	10			
3	16			
4	40			
5	50			
6	46			
7	10			
8	12			
9	4			
10	2			

Calcular el percentil que li correspon a una puntuació directa de 4 i a una de 7

Puntuacions Típiques

$$z = \frac{x - \bar{x}}{S_x}$$

x - PD del subjecte

\bar{x} - Mitjana del grup normatiu (GN)

S_x - desviació típica del GN

Si $z \sim N(0, 1) \Rightarrow -\infty \div \infty \Rightarrow -2.58 \div 2.58 \rightarrow 99\%$

Mesura quant se separa cada subj. de la mitjana \rightarrow
Distància en unitats de desviació tipus (unitat de mesura constant).

Inconvenients: - Valors negatius
- Valors decimals

Exemple: Puntuacions Típiques

Un test aplicat a un grup de 1500 subjectes es va distribuir normalment amb una mitjana de 25 i una variança de 16. Si un subjecte obté en aquest mateix test una puntuació directa de 31, quina puntuació típica li correspon?

$$z = \frac{31 - 25}{4} = 1.5$$

Exemple: Puntuacions Típiques

En una prova d'aritmètica (amb mitjana 20 i desviació típica 4) en Joan obté una puntuació de 24 i en Josep una de 28. Aquestes dues puntuacions no ens indiquen directament el grau de coneixements aritmètics dels dos subjectes, ni ens permeten saber a quina distància estan entre sí.

Si calculem la puntuació típica respectiva (mitjançant la fórmula de z), veurem que en Joan obté una z d'1 i en Josep una de 2.

Joan	—————→	$x_i = 24$	$Z = 1$
Josep	—————→	$x_i = 28$	$Z = 2$

Exercici: Puntuacions Típiques

Si a més fan una altra prova de trigonometria amb $\bar{x} = 35$ i $S_x = 6$, i obtenen les següents puntuacions

Joan \longrightarrow $x_i = 40$ $Z = 0,8\bar{3}$
Josep \longrightarrow $x_i = 38$ $Z = 0,5$

Globalment, quin dels dos té més coneixements de matemàtiques?

Exercici: Percentils i Puntuacions Típiques

Disposem de la següent taula de dades, amb la columna x de puntuacions obtingudes en una prova i la columna f_i (freqüència absoluta) amb el nombre de participants que han obtingut cadascuna de les possibles puntuacions:

x	f_i	F_i	P	z
0	4			
1	15			
2	13			
3	24			
4	63			
5	81			
6	93			
7	85			
8	99			
9	108			
10	75			
11	31			
12	29			
13	41			
14	27			
15	12			

Obtenir, per cadascuna de les puntuacions directes, els percentils i les puntuacions típiques, sabent que:

$$\bar{x} = 7.77125$$

$$S_x = 3.1518$$

Puntuacions Típiques Normalitzades

Assumim que les característiques mesurades per un test es distribueixen normalment en la població

Al treballar amb mostres, les dades generalment no són perfectament normals

Puntuacions típiques normalitzades (z_n)

PD → Percentils → Puntuacions Típiques Normalitzades

Si la distribució de les dades fos perfectament normal, aleshores $z = z_n$

Exemple: Puntuacions típiques i puntuacions típiques normalitzades

x	f_i	P	z	P	p	z_n
0	4	0.25	-2.4656	1	0,01	-2.3263
1	15	1.4375	-2.1484	1	0,01	-2.3263
2	13	3.1875	-1.8311	3	0,03	-1.8808
3	24	5.5	-1.5138	6	0,06	-1.5548
4	63	10.9375	-1.1965	11	0,11	-1.2265
5	81	19.9375	-0.8792	20	0,20	-0.8416
6	93	30.8125	-0.5619	31	0,31	-0.4959
7	85	41.9375	-0.2447	42	0,42	-0.2019
8	99	53.4375	0.0726	53	0,53	0.0753
9	108	66.375	0.3898	66	0,66	0.4125
10	75	77.8125	0.7071	78	0,78	0.7722
11	31	84.4375	1.0244	84	0,84	0.9945
12	29	88.1875	1.3417	88	0,88	1.1750
13	41	92.5625	1.6589	93	0,93	1.4758
14	27	96.8125	1.9763	97	0,97	1.8808
15	12	99.25	2.2935	99	0,99	2.3263

$$\bar{x} = 7.77125$$

$$S_x = 3.1518$$

veure taula distribució normal z

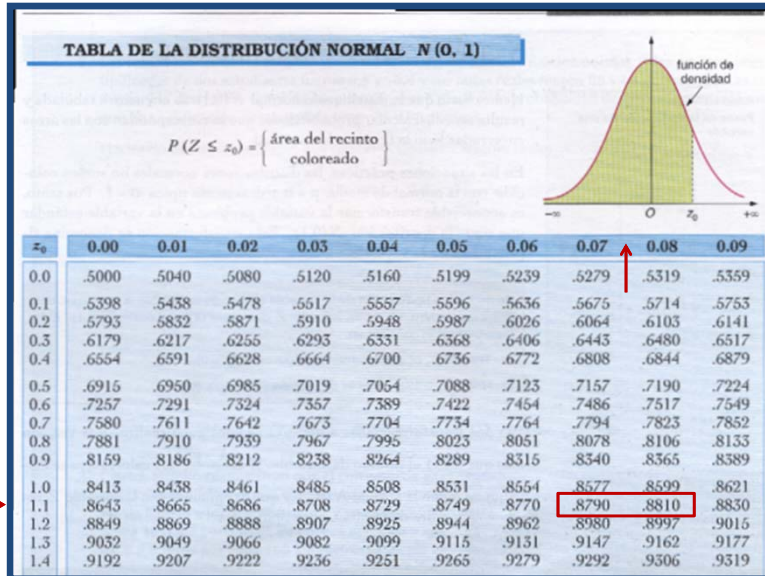


Figura extreta de <http://ieszaframagon.com>

Correspondència entre Puntuacions típiques i Percentils

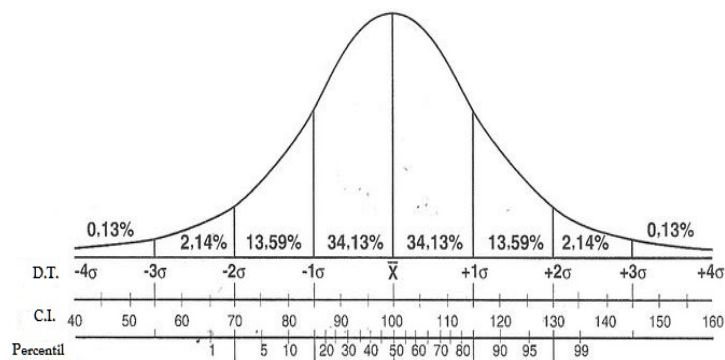


Figura extreta de <http://psicopsi.com>

Puntuacions Típiques Derivades

$$z_d = \bar{x}_d + s_d \cdot z_x$$

On: \bar{x}_d - mitjana de les puntuacions típiques derivades.
 s_d - desviació de les puntuacions típiques derivades.
 z_x - puntuació típica original

Puntuacions Típiques Derivades Normalitzades

En cas que la puntuació derivada s'obtingui de la z_n ,
parlarem de **Puntuacions típiques derivades
normalitzades**

$$z_{d_n} = \bar{x}_d + s_d \cdot z_n$$

On: \bar{x}_d - mitjana de les puntuacions típiques derivades.
 s_d - desviació de les puntuacions típiques derivades.
 z_n - puntuació típica normalitzada original

Tipus de Puntuacions Típiques Derivades

A cada z li correspon una puntuació derivada:

$$\text{Puntuacions T} \rightarrow T = 50 + 10 z$$

$$\text{Puntuacions D} \rightarrow D = 50 + 20 z$$

$$\text{QI} \rightarrow \text{QI} = 100 + 15 z$$

A diferents valors de z els correspon la mateixa puntuació derivada:

$$\text{Sten o decatipus (1-10)} \rightarrow \text{Sten} = 5.5 + 2 z$$

$$\text{Estanines (1-9)} \rightarrow e = 5 + 2 z$$

$$\text{Pentas (1-5)} \rightarrow p = 3 + z$$

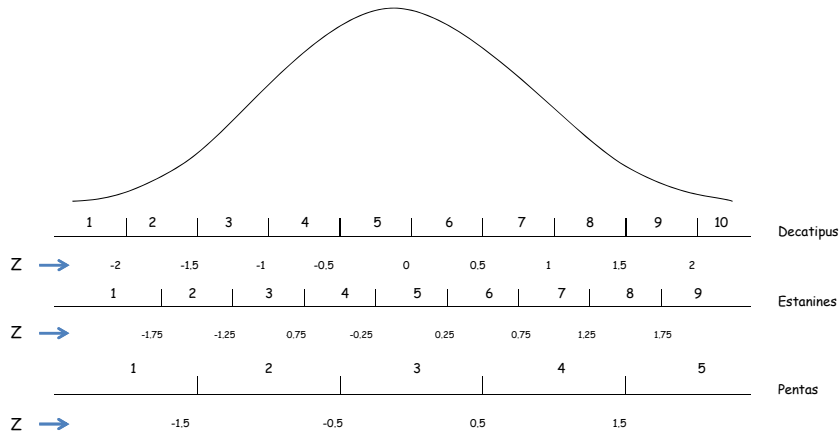
Exemple: Puntuacions Típiques Derivades

	z	T	D	Sten	Estanina
0	-2,47	25	1	1	1
1	-2,15	29	7	1	1
2	-1,83	32	13	2	1
3	-1,51	35	20	2	2
4	-1,20	38	26	3	3

$$T = 50 + 10 (-2,47) = 25,3 \rightarrow 25$$

$$\text{Sten} = 5,5 + 2 (-2,47) = 0,56 \rightarrow 1$$

Puntuacions Típiques Derivades



Interpretació Puntuacions Típiques i Típiques Derivades

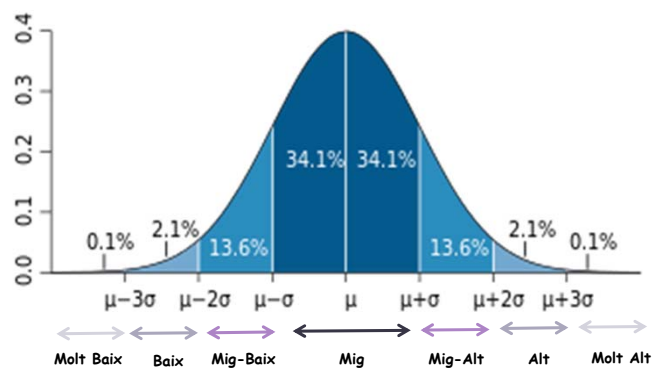


Figura modificada de https://es.wikipedia.org/wiki/Distribuci3n_normal

Exemple: Trobar el valor en PTD per la PD 9 de l'exemple anterior

$$z = 0.3898$$

$$\text{Puntuacions } T \rightarrow T = 50 + 10 \cdot 0.3898 = 53.898 \rightarrow 54$$

$$\text{Puntuacions } D \rightarrow D = 50 + 20 \cdot 0.3898 = 57.796 \rightarrow 58$$

$$QI \rightarrow QI = 100 + 15 \cdot 0.3898 = 105.847 \rightarrow 106$$

$$\text{Decatipus (1-10)} \rightarrow \text{Sten} = 5.5 + 2 \cdot 0.3898 = 6.28 \rightarrow 6$$

$$\text{Estanines (1-9)} \rightarrow e = 5 + 2 \cdot 0.3898 = 5.7796 \rightarrow 6$$

$$\text{Pentas (1-5)} \rightarrow p = 3 + 0.3898 = 3.3898 \rightarrow 3$$

Exercici: Realitza la transformació de puntuacions de les següents dades

x	f_i	F_i	P	z	T	D	QI	$Sten$	$Estanines$	$Pentas$
0	4									
1	15									
2	13									
3	24									
4	63									
5	81									
6	93									
7	85									
8	99									
9	108									
10	75									
11	31									
12	29									
13	41									
14	27									
15	12									

$$\bar{x} = 7.77$$

$$S_x = 3.15$$

Exercici: Baremació

Un individu obté una puntuació de 70 en una prova de mitjana 65, desviació típica 9 i una fiabilitat de 0,85.

- 1) Quina seria la seva puntuació en una escala T?
- 2) Quin seria l'interval (en puntuacions T) de la puntuació vertadera d'aquest individu (IC=95%, $z=1.96$)?
- 3) Quin és el nivell en el que es trobaria l'individu?

6. Manual del Test

TEMA: Manual del test

Lectura:

Muñiz, J., y Fernández-Hermida, J.R. (2010). La opinión de los psicólogos españoles sobre el uso de los tests. *Papeles del Psicólogo*, 31(1), 108-121.

El manual com a guia de l'usuari

1. Constructe que cal mesurar i teoria subjacent - Que es vol mesurar?
2. Creació d'ítems - Indicadors observables de la conducta externa del subjecte
3. Recollida de dades
4. Anàlisi d'ítems: ID, Discriminació, Distractors, etc.
5. Selecció d'ítems: Complir els requisits necessaris
6. Obtenció del test normatiu: Fiabilitat i Validesa
7. Selecció del grup normatiu
8. Transformació de les puntuacions

Fonts més habituals d'error en l'adaptació de tests

Context d'utilització – sociocultural

Construcció de la prova – traducció inversa (comprovar
fiabilitat, validesa i estandardització)

Aplicació

Interpretació dels resultats – equivalència entre versions

Aspectes ètics i deontològics: APA

Competència

Integritat

Responsabilitat professional i científica

Respecte pels drets i la dignitat de les persones

Procurar el benestar de les persones

Responsabilitat social



Utilitat avaluació

Possibilitat de realització

Adequació

Precisió

Utilització dels tests

Muñiz, J. (2005). Utilización de los tests. En J. Muñiz, A.M. Fidalgo, E. García-Cueto, R. Martínez y R. Moreno (Eds.). *Análisis de los ítems*, (pp. 133-172). Madrid: La Muralla, S.A.

Avaluació de la qualitat del test

Descripció del test

Valoració de les característiques del test

Valoració global del test

Directrius elaborades per la Comissió Internacional de Tests (ITC) per l'ús adequat dels test (Bartram, 2001) - Objectiu: millorar l'ús dels tests

<http://www.efpa.be>

<http://www.cop.es>

<http://www.intestcom.org>

5

Índex Directrius Internacionals per a l'ús dels Tests (Colegio Oficial de Psicólogos - Comisión Internacional de Tests)

Introducción

Directrices

Objetivo

Ámbito de aplicación

Destinatarios

Aspectos contextuales

Conocimientos

1. Uso ético de los tests

- 1.1 Actuar de forma ética y profesional.
- 1.2 Asegurarse de que son competentes para el uso de los tests.
- 1.3 Responsabilizarse del uso que hacen de los tests.
- 1.4 Asegurarse de que los materiales del test están seguros.
- 1.5 Asegurarse de que los resultados de los tests se tratan confidencialmente.

2. Utilización adecuada de los tests

- 2.1 Estimar la utilidad potencial de los tests en una situación evaluativa.
- 2.2 Elegir tests técnicamente correctos y adecuados a cada situación.
- 2.3 Prestar atención a los aspectos relacionados con el sesgo de los tests.
- 2.4 Hacer los preparativos necesarios para la aplicación del test.
- 2.5 Aplicar los tests adecuadamente.
- 2.6 Puntuar y analizar los resultados de los tests con precisión.
- 2.7 Interpretar los resultados adecuadamente.
- 2.8 Comunicar los resultados de forma clara y precisa.
- 2.9 Revisión de la adecuación del test y de su uso.

Colegio Oficial de Psicólogos Comisión - Internacional de Tests

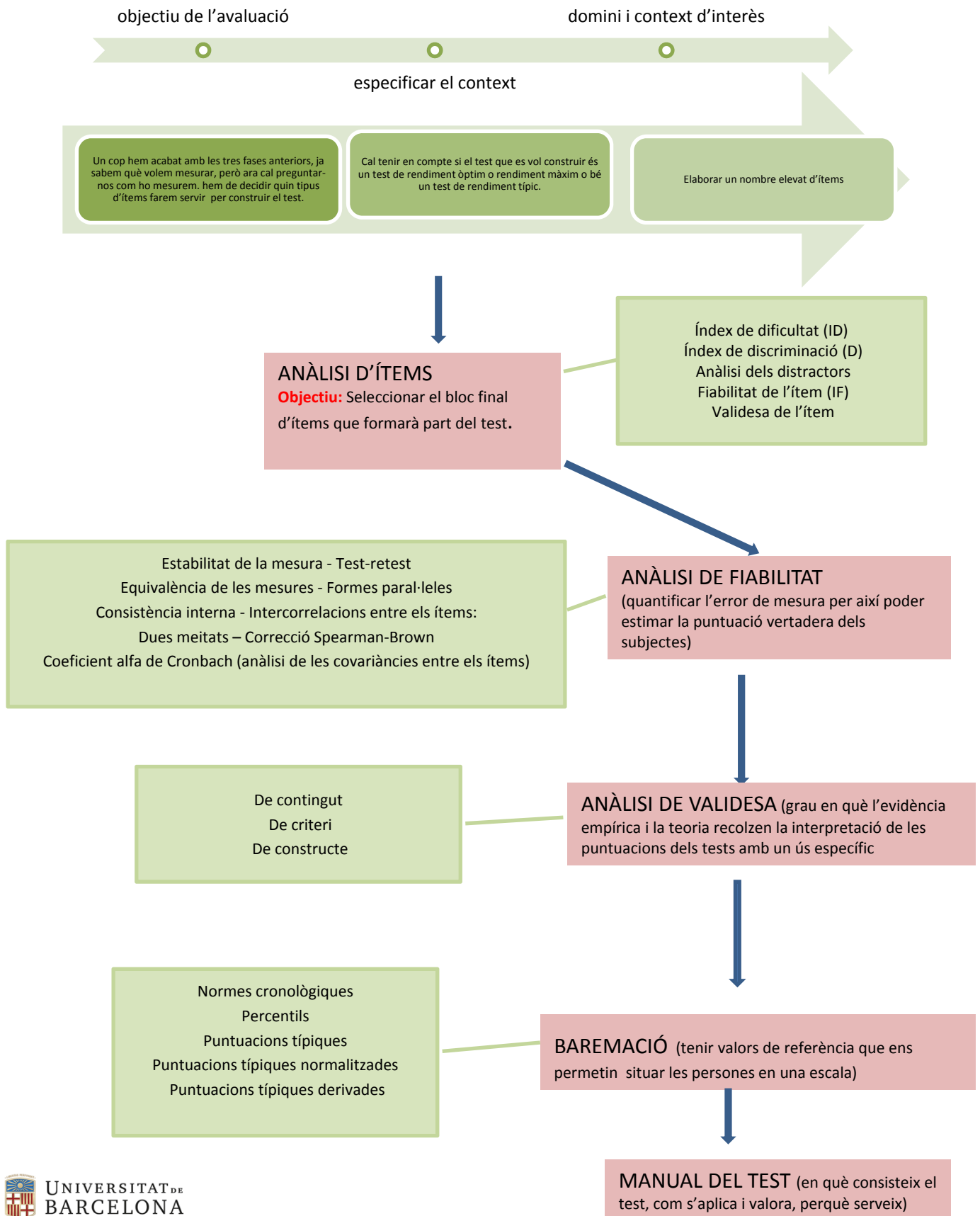
2.2. Elegir tests técnicamente correctos y adecuados a cada situación

Los usuarios competentes deberían:

- 2.2.1. Examinar toda la información disponible sobre los tests potencialmente adecuados antes de elegir un test concreto.
- 2.2.2. Comprobar que la documentación técnica sobre el test proporciona suficiente información para evaluar los siguientes aspectos:
 - a. Amplitud y representatividad del contenido del test, adecuación de los grupos normativos utilizados, nivel de dificultad de los contenidos, etc.
 - b. Precisión de la medición y fiabilidad para las poblaciones pertinentes.
 - c. Validez para las poblaciones pertinentes y su aplicabilidad para el uso que se hace del test.
 - d. Ausencia de sesgo para los grupos con los que se utilizará .
 - e. Aceptación por parte de quienes están implicados en su uso, incluyendo la pertinencia y validez aparente percibidas.
 - f. Aspectos prácticos, tales como tiempo requerido, coste, o recursos que se necesitan.
- 2.2.3. Evitar el uso de tests que tengan una documentación técnica inadecuada o poco clara.
- 2.2.4. Utilizar tests sólo para aquellos objetivos para los cuales se dispone de una validez empírica adecuada y pertinente.
- 2.2.5. No aceptar un test basándose únicamente en su validez aparente, recomendaciones de otros usuarios, o consejos de quienes tienen intereses comerciales.
- 2.2.6. Responder a las preguntas de las personas implicadas (personas evaluadas, padres, supervisores, representantes legales, etc.), dándoles suficiente información para que entiendan por qué se eligió el test.

7. Annexes

CONSTRUCCIÓ DE TESTS



FORMULARI

$$\bar{x} = \frac{\sum x_j}{N} = \sum ID_j$$

$$ID = \frac{A}{N}$$

$$Pc = A - E / (k - 1)$$

$$ID = \frac{A - E / (k - 1)}{N}$$

$$\Phi = \frac{p_{xy} - p_x p_y}{\sqrt{p_x q_x p_y q_y}}$$

$$r_{bp} = \frac{\bar{x}_p - \bar{x}_x}{S_x} \sqrt{p/q}$$

$$D = \frac{n_s}{N_s} - \frac{n_i}{N_i}$$

$$r_{xx} = \frac{2 r_{x_1 x_2}}{1 + r_{x_1 x_2}}$$

$$S_x = \sum S_j \cdot r_{jx}$$

$$S_j^2 = p \cdot q$$

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y}$$

$$r_b = \frac{\bar{x}_p - \bar{x}_x}{S_x} \frac{p}{y}$$

$$r_t = \sqrt{\frac{(ad) - (bc)}{n^2 y y'}}$$

$$Cr = \frac{r_{jx} S_x - S_j}{\sqrt{S_x^2 + S_j^2 - 2 r_{jx} S_x S_j}}$$

$$IF = S_j D_j$$

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum S_j^2}{S_x^2} \right)$$

$$r_{xx} = \frac{S_v^2}{S_x^2}$$

$$\alpha = KR20 = \frac{n}{n-1} \left(1 - \frac{\sum p_j \cdot q_j}{S_x^2} \right)$$

$$k = \frac{n_i + n_a}{n_i}$$

$$R_{xx} = \frac{k r_{xx}}{1 + (k-1) r_{xx}}$$

$$k = \frac{R_{xx} (1 - r_{xx})}{r_{xx} (1 - R_{xx})}$$

$$V_{int} = x_i \pm z \cdot S_e$$

$$S_e = S_x \sqrt{1 - r_{xx}}$$

$$y_i = r_{xy} \frac{S_y}{S_x} (x_i - \bar{x}) + \bar{y}$$

$$Y_{int} = y_i \pm z \cdot S_{xy}$$

$$S_{xy} = S_y \sqrt{1 - r_{xy}^2}$$

$$P_c = \frac{\text{Vertaders positius } (f_{11}) + \text{Vertaders negatius } (f_{22})}{N}$$

$$\text{Sensibilitat} = \frac{\text{Vertaders positius } (f_{11})}{\text{Total persones amb trastorn } (f_{.1})}$$

$$\text{Especificitat} = \frac{\text{Vertaders negatius } (f_{22})}{\text{Total persones sanes } (f_{.2})}$$

$$K = \frac{F_c - F_a}{N - F_a}$$

$$F_c = \text{Vertaders positius } (f_{11}) + \text{Vertaders negatius } (f_{22})$$

$$F_a = \frac{f_{1.} \cdot f_{.1} + f_{2.} \cdot f_{.2}}{N}$$

$$S_{e(K)} = \sqrt{\frac{F_a}{N \cdot (N - F_a)}}$$

$$IC \Rightarrow K \pm z \cdot S_{e(K)}$$

$$QI = \frac{EM}{EC} \cdot 100$$

$$P = \frac{F_{i-1} + 0.5 f_i}{N} \cdot 100$$

$$z = \frac{x - \bar{x}}{S_x}$$

$$z_d = \bar{x}_d + s_d \cdot z_x$$

$$z_{d_n} = \bar{x}_d + s_d \cdot z_n$$

$$T = 50 + 10 \cdot z$$

$$D = 50 + 20 \cdot z$$

$$QI = 100 + 15 \cdot z$$

$$\text{Sten} = 5.5 + 2 \cdot z$$

$$e = 5 + 2 \cdot z$$

$$p = 3 + z$$