

## QUARTA ACTIVITAT D'AVALUACIÓ CONTINUADA

Gemma Pérez Moliner

### 1.1.

Els professors de cinquè curs de primària d'una escola sospiten que a la seva aula es dona un cas d'assetjament entre escolars ("bullying"). Donada la sensibilització que s'ha produït en els darrers temps (degut a notícies similars que han aparegut en premsa) la preocupació dels professors els porta a voler comprovar si en la seva aula es dona aquesta situació i a identificar als protagonistes (assetjats i assetjadors) per a posar-hi remei abans que la situació sigui més greu.

El consell de professors posa aquesta tasca en mans del psicòleg del centre.

### 1.2.

El psicòleg decideix administrar un test a tota l'aula que avaluï l'agressivitat entre escolars i n'identifiqui els protagonistes. Es decideix pel test "Bull-s – Medida de la agresividad entre escolares" de l'Editorial Albor-Cohs (2000).

La Forma A del test s'administra col·lectivament als alumnes (d'entre 7 i 16 anys). La Forma B, de manera individual als professors (d'Educació Primària i Secundària). El temps d'administració varia entre 25 i 30 minuts.

Aquest test està indicat per a l'anàlisi de les característiques socio-afectives en el grup d'iguals. Facilita la detecció dels individus implicats en les situacions d'assetjament (víctimes o assetjats, i "bullies" o assetjadors), i permet l'anàlisi dels aspectes situacionals en les relacions d'agressivitat entre companys, així com del clima social de l'aula.

Els resultats que s'obtenen amb aquest test consisteixen en: Sociograma del grup, Posició sociomètrica individual, Punts de tall per a valors significatius en les variables d'agressió i de victimització i, Freqüències en aspectes situacionals.

Aquests resultats obtinguts després de l'administració i anàlisi, permetran al psicòleg identificar els assetjats i els assetjadors, de manera que podrà prestar una atenció més personalitzada a aquests alumnes, al mateix temps que li permetrà un seguiment individualitzat de la seva evolució. També podrà realitzar un seguiment general de l'ambient social dins el grup.

### 1.3.

El psicòleg podrà inferir quin és el clima social que impera a l'aula i el nivell de gravetat que li atribueixen els propis escolars. Així mateix, podrà identificar aquells alumnes que actuen com a "matons" i aquells a qui es dirigeixen les agressions, permetent-li elaborar un perfil de característiques psico-socials d'aquests nens i obtenir informació sobre les característiques de l'agressió.

Donat que el test ha estat ja publicat i validat pels seus autors, hem d'entendre que les característiques tècniques del mateix ja han estat suficientment provades i validades; per tant, pel que fa al procés de resposta, l'estructura interna i les conseqüències de l'avaluació, podem suposar que s'adeqüen als objectius que pretén el test. Aquestes característiques haurien d'estar explicitades en el manual corresponent.

No obstant, el nostre psicòleg voldrà comprovar que el test que pretén utilitzar és adient per a l'avaluació que durà a terme, és a dir, que és vàlid per avaluar la problemàtica que se li presenta.

Validesa de contingut:

Així, la única estratègia que li quedaria al nostre psicòleg seria provar la relació entre el contingut del test i el constructe que pretén de mesurar. És a dir, si el contingut del test és vàlid per a mesurar la situació social a l'aula i identificar els alumnes implicats.

L'evidència relativa al contingut implica que els ítems del test siguin representatius i rellevants del constructe. Per a obtenir aquesta evidència, el psicòleg haurà de definir prèviament quina és la situació que vol mesurar, quines són les característiques de la mateixa i en què consisteixen les actituds i comportaments que la representen (per exemple, a partir de les observacions dels professors que li han encarregat l'avaluació, la literatura i estudis existents sobre el tema, i les dades conegudes d'altres casos). Després, com a expert, comprovarà que el test representa tots els aspectes d'aquesta definició i, que no es centra en altres aspectes aliens a la mateixa.

Validesa concurrent:

Per una altra banda, una estratègia diferent per a comprovar l'adequació del test a l'objectiu perseguit, consistiria a acompanyar la passació del mateix amb entrevistes amb els professors que han detectat el problema i amb els alumnes de classe.

Així, a partir de les dades obtingudes en les entrevistes (sobre l'ambient de l'aula, les implicacions dels diferents alumnes, etc.) i mitjançant la seva correcta codificació i comparació, podrà contrastar si ambdues mesures (test i entrevista) estan relacionades i per tant, la pertinència dels resultats obtinguts en el test.

I encara, podria passar el mateix test als alumnes d'una altra aula (amb les mateixes característiques socio-culturals) on tots els indicis apuntessin a que no existeix la problemàtica descrita pels professors. Els resultats obtinguts en l'altra aula haurien de coincidir amb les observacions dels professors i ser diferents dels obtinguts en la nostra aula.

### 2.1.

Sabem que en un ítem dicotòmic, la proporció de respostes codificades com a 1 ( $p_j$ ) multiplicada per la proporció de respostes codificades com a 0 ( $q_j$ ), és igual a la variància d'aquest ítem.

Com que tenim el nombre de respostes donades a cada un dels ítems (10, una per subjecte), podem calcular tots els productes possibles i, per tant, tots els valors de variància possibles en aquest test:

	$p_j$	$q_j$	$p_j \times q_j$
0 respostes 1 + 10 respostes 0 (o viceversa)	0	1	0
1 resposta 1 + 9 respostes 0 (o viceversa)	0,1	0,9	0,09
2 respostes 1 + 8 respostes 0 (o viceversa)	0,2	0,8	0,16
3 respostes 1 + 7 respostes 0 (o viceversa)	0,3	0,7	0,21
4 respostes 1 + 6 respostes 0 (o viceversa)	0,4	0,6	0,24
5 respostes 1 + 5 respostes 0 (o viceversa)	0,5	0,5	0,25

Llavors, tenint en compte que les variàncies possibles dels ítems en aquest test són iguals als valors de la columna " $p_j \times q_j$ ", veiem que només una de les variàncies donades (0,15 per l'ítem 1) no s'ajusta a cap de les variàncies possibles. Aquest seria el valor erroni.

I el valor correcte més proper a aquest valor erroni seria 0,16 .

### 2.2.

L'ítem que menys contribueix a la consistència interna d'aquest test és l'ítem 2.

Podem entendre que (per la seva similitud), els valors de "Correlació Ítem-PT" corresponen als valors de "correlació de l'ítem amb el total corregit", és a dir, a la correlació de la puntuació de cada ítem amb la puntuació total del test calculada com a suma de tots els ítems excepte l'avaluat.

El valor de la correlació de l'ítem 2 amb el total corregit és la més baixa de totes (0,30), i, tenint en compte que la correlació de l'ítem amb el total corregit és un indicador directe del grau de contribució d'un ítem a la consistència interna del test; podem dir que la correlació més baixa correspon a la contribució més baixa.

### 3.

La resposta correcta és la c): menys de 20.

Sabem que, en general, la fiabilitat augmenta en relació directa amb el nombre de components de la mesura, és a dir, a major nombre d'ítems, més fiabilitat.

Així, en afegir 5 ítems paral·lels (això és, que cada ítem aporta la mateixa quantitat de fiabilitat), la fiabilitat del test augmentarà.

També sabem que la fiabilitat d'un instrument està inversament relacionada amb l'error típic de mesura, o el que és el mateix, com més fiable sigui l'instrument més s'acostaran les puntuacions observades a les vertaderes, la qual cosa ens indicarà que s'ha comès menys quantitat d'error en la mesura.

Així, al augmentar la fiabilitat, l'error típic de mesura disminuirà i, per tant, la variància dels errors serà més petita.

### 4.

#### 1)

Suposem que tenim un test amb uns paràmetres determinats, una fiabilitat concreta i un cert error aleatori de mesura. Com hem vista abans, una manera de disminuir el error de mesura i augmentar-ne la fiabilitat seria afegint un cert nombre d'ítems paral·lels al nostre test.

Per una altra banda, sabem que la fiabilitat del test afecta i limita la validesa, de manera que la falta de fiabilitat fa que disminueixi la validesa.

Així, afegint ítems paral·lels al nostre test, a part d'augmentar la fiabilitat, haurem millorat la validesa.

2)

La validesa global del test ve determinada per la validesa de cada un dels ítems que el componen. Partint d'un test donat que hauríem construït, podríem sotmetre tots i cada un dels ítems del test a una anàlisi de contingut per part d'experts independents.

Donada una definició clara del constructe que pretenem de mesurar amb el test, els experts avaluarien el contingut dels ítems plantejats i opinarien si aquests són rellevants pel constructe o no.

El fet d'eliminar aquells ítems que fossin considerats unànimement com a no rellevants pels experts, faria millorar la validesa del nostre test.

3)

Donada una definició clara i precisa del constructe que es pretén de mesurar, es pot avaluar la representativitat del conjunt d'ítems.

Altre cop, podríem ajudar-nos de l'opinió d'experts independents perquè avaluessin si els ítems del nostre test abarquen tot el domini del constructe. Si no és així, si els ítems no fossin prou representatius del constructe, el fet d'afegir ítems no redundants al nostre test faria augmentar la validesa.

4)

Atenent als processos de resposta també es pot millorar la validesa del test.

Si coneixem i controlem els factors que incideixen en els processos de resposta evitarem que els subjectes que responen el nostre test donin respostes irrelevantes. Si es donessin aquestes respostes irrelevantes estaríem introduint una font d'error al mesurar allò que volem mesurar, és a dir, que les respostes no informarien sobre el constructe i, per tant, que el test perdria validesa.

Les maneres de controlar els processos de resposta passen per construir ítems no ambigus, amb un contingut adequat i que siguin comprensible per als avaluats; aconseguir que el test presenti validesa aparent; elaborar unes instruccions clares i comprensibles, i; dissenyar l'estructura del test de manera que no permeti biaixos a l'hora de triar una resposta (per exemple: sempre la mateixa opció).

5.

Amb les dades:

Puntuació mitjana del grup: 2,36

Desviació estàndard de les puntuacions compostes: 1,96

Càlcul de les puntuacions z:  $z_{xi} = \frac{x_i - \mu_x}{\sigma_x}$

Càlcul de les puntuacions T:  $T_i = 50 + 10z_{xi}$

Càlcul dels percentils associats a cada una de les puntuacions:  $P_{ci} = \frac{f_{aa} + 0,5f_i}{N} \times 100$

Càlcul dels Enneatipus: Enneatipus = 5 + 2z<sub>n</sub>

Càlcul dels Decatipus: Decatipus = 5,5 + 2z<sub>n</sub>

Obtenim:

	<b>z</b>	<b>T</b>	<b>f<sub>i</sub></b>	<b>F<sub>a</sub></b>	<b>Pc</b>	<b>z<sub>n</sub></b>	<b>Enneatipus</b>	<b>Decatipus</b>
Puntuació directa 0	-1,2	38	118	118	11,87	-1,18	2,64	3,14
Puntuació directa 1	-0,69	43	87	205	32,49	-0,45	4,09	4,59
Puntuació directa 2	-0,18	48	69	274	48,19	-0,05	4,91	5,41
Puntuació directa 3	0,33	53	78	352	62,98	0,33	5,66	6,16
Puntuació directa 4	0,83	58	51	403	75,96	0,70	6,41	6,91
Puntuació directa 5	1,34	63	53	456	86,42	1,10	7,20	7,70
Puntuació directa 6	1,85	69	41	497	95,88	1,74	8,47	8,97

Per la puntuació directa 6:

$$z_{xi} = \frac{x_i - \mu_x}{\sigma_x} = \frac{6 - 2,36}{1,96} = 1,85$$

$$T_i = 50 + 10z_{xi} = 50 + (10 \times 1,85) = 68,5 \approx 69$$

f<sub>i</sub>: utilitzant la fórmula FRECUENCIA d'Excel

$$P_{ci} = \frac{f_{aa} + 0,5f_i}{N} \times 100 = \frac{456 + (0,5 \times 41)}{497} \times 100 = 95,88$$

$Z_n$ : Utilitzant la fórmula DISTR.NORM.ESTAND.INV d'Excel

Enneatipus =  $5 + 2z_n = 5 + (2 \times 1,74) = 8,47$

Decatipus =  $5,5 + 2z_n = 5,5 + (2 \times 1,74) = 8,97$

6.

L'estudi s'ha realitzat amb una mostra de 497 subjectes, dels quals 210 homes (42,25%) i 287 dones (57,75%).

La mostra s'ha obtingut entre els alumnes de psicometria del curs Q1-2005-06, els quals, a la vegada passaven el test a dues persones més del seu entorn.

Les edats dels subjectes estaven compreses entre 13 i 76 anys, essent la mitja 36,25 anys, la mediana 35 anys i la moda 32 anys. La desviació estàndard de les edats ha sigut de 10,61.

Pel que fa als homes, les edats estaven compreses entre els 13 i els 68 anys, essent la mitja de 36,29 anys, la mediana 34 anys i la moda 34 anys. La desviació estàndard de les edats ha sigut de 9,91.

Les edats de les dones estaven compreses entre els 16 i els 76 anys, essent la mitja de 36,22 anys, la mediana 35 anys i la moda 32 anys. La desviació estàndard de les edats ha sigut de 11,11.

El test aplicat, EPQ-RA, està format per 4 escales (Extraversió, Neuroticisme, Psicoticisme i una escala de Dissimul/Conformitat) amb 6 ítems cada una.

Les puntuacions compostes obtingudes per a cada escala, es poden resumir com a:

	E	N	P	L
Mitjana	3,36	2,21	2,22	2,34
Desviació estàndard	1,74	1,88	1,48	1,74

En general, podem dir que els subjectes puntuen més alt en l'escala E que en qualsevol de les altres i que la variabilitat més gran la presenta l'escala N, així com la més petita la presenta l'escala P.

Si obtenim la matriu de correlacions entre les quatre escales:

	E	N	P	L
E				
N	-0,107			
P	0,187	0,263		
L	-0,081	-0,050	-0,161	

veiem que els valors de les correlacions són molt pròxims a 0, la qual cosa ens indicaria que les quatre escales són molt independents les unes de les altres. Observem que la correlació entre les escales N i P presenten un valor de correlació més elevat (0,263), encara que no és un valor significativament alt, això podria indicar que els constructes als que representen aquestes escales podrien presentar característiques comunes.

De la correlació entre la puntuació composta de cada escala i els ítems que la formen hem obtingut les saturacions factorials; i hem vist que tots els ítems presenten saturacions factorials elevades excepte l'ítem 11 de l'escala E (0,499), els ítems 5 (0,459), 17 (0,468) i 21 (0,460) de l'escala P i l'ítem 16 de l'escala N (0,473). No obstant, aquestes correlacions no són baixes sinó moderades. Així, podem concloure que les puntuacions compostes de cada escala representen un bon resum dels ítems que les componen.

També, de les correlacions ítem-ítem de cada escala, extraïem que, en general aquestes són baixes (exceptuant les correlacions entre els ítems 7-14 de l'escala E, i; 15-19, 13-22, 12-15, 13-15 i 12-19 de l'escala N, en que les correlacions són moderades).

En general podem dir que les quatre escales són unidimensionals.

Aplicant la fórmula del coeficient alfa per a ítems dicotòmics, obtenim la fiabilitat de cada una de les escales:

	E	N	P	L
alfa	0,661	0,760	0,425	0,630

aquests resultats ens mostren que en les escales E, N i L, el resultat és més proper a 1 que a 0, sense arribar a ser 1, per tant, podem considerar que els ítems que formen aquestes escales estan ben relacionats entre ells sense ser redundants, i, per tant, que existeix consistència interna; en canvi, l'escala P ens dona un valor d'alfa moderat, la qual cosa voldria dir que els ítems de l'escala P no representen tan bé el seu constructe com ho fan els ítems de les altres escales.

Per una banda hem comprovat que les quatre escales són unidimensionals i que les puntuacions compostes representen bé els ítems. Per altra banda, hem vist que (tot i ser una adaptació) les escales presenten uns nivells de fiabilitat acceptables.

Per tant, si que podem dir que el test EPQ-RA compleix els requisits de Validesa i Fiabilitat que tot test ha de complir.

\*\*\*\*\*