

**COLEGIUL PSIHOLOGILOR DIN ROMÂNIA
COMITETUL DIRECTOR**

**NORMELE DE AVIZARE A METODELOR ȘI TEHNICILOR DE EVALUARE ȘI
ASISTENȚĂ PSIHOLOGICĂ**

**Capitolul I
Dispoziții generale**

Art. 1 (1) Potrivit art. 33 lit. h) din Legea nr. 213/2004 privind exercitarea profesiei de psiholog cu drept de liberă practică, înființarea, organizarea și funcționarea Colegiului Psihologilor din România, publicată în M.O., Partea I, nr. 492 din 1 iunie 2004, denumită în continuare *lege*, *Comitetul director* al *Colegiului Psihologilor din România*, denumit în continuare *Comitetul director*, elaborează normele de avizare a metodelor și tehnicilor de evaluare și asistență psihologică utilizate de către psihologii cu drept de liberă practică din România.

(2) *Comisia metodologică* constituită în cadrul *Comitetului director* reprezintă unica autoritate, care, potrivit art. 36 lit. b) din *lege*, analizează și apreciază metodele și tehnicile de evaluare și asistență psihologică.

(3) În urma analizării metodelor și tehnicilor de evaluare și asistență psihologică, *Comisia metodologică* propune *Comitetului director* eliberarea sau neeliberarea *avizului* pentru metodele și tehnicile de evaluare și asistență psihologică analizate, precum și *modul de utilizare* al acestora.

(4) *Comitetul director* eliberează *avizul* pentru metodele și tehnicile de evaluare și asistență psihologică în baza propunerii *Comisiei metodologice* sau o adresă în care sunt menționate motivele care au stat la baza neacordării avizului.

Art. 2 În vederea instituirii și promovării standardelor de calitate a serviciilor psihologice prestate de către psihologul cu drept de liberă practică, *Comitetul director* stabilește prin prezentele *norme* condițiile acordării avizului pentru metodele și tehnicile de evaluare și asistență psihologică utilizate de către psihologii cu drept de liberă practică din România.

Art. 3 (1) Avizele eliberate de către *Comitetul director* sunt solicitate în mod obligatoriu de către *producătorii sau distribuitorii de metode și tehnici de evaluare și asistență psihologică*, persoane fizice sau juridice, române sau străine, care intenționează să le introducă pe piața românească în vederea comercializării sau utilizării cu titlu gratuit.

(2) În desfășurarea activităților profesionale prevăzute de art. 5 din *lege*, psihologii cu drept de liberă practică utilizează numai metodele și tehnicile de evaluare și asistență psihologică avizate de către *Comitetul director*.

(3) *Evidența* metodelor și tehnicilor de evaluare și asistență psihologică este ținută și actualizată de către *Colegiul Psihologilor din România*, prin înscrierea automată a acestora în urma obținerii avizului în *Catalogul metodelor și tehnicilor de evaluare și asistență psihologică din România*.

(4) *Catalogul metodelor și tehnicilor de evaluare și asistență psihologică* este un document public, care este publicat de către *Comitetul director* prin editare anuală.

Capitolul II

Definiții și precizări terminologice

Art. 4 (1) *Normele de avizare a metodelor și tehnicilor de evaluare și asistență psihologică* constau într-un set minim de prescripții metodologice, explicit formulate, care, adoptate și respectate atât de către *producătorii și distribuitorii autorizați* cât și de către psihologii cu drept de liberă practică, asigură coerență în formularea investigațiilor și a concluziilor psiho-diagnostice.

(2) *Metodele și tehnicile de evaluare psihologică* constau în totalitatea demersurilor fundamentate științific, efectuate în scopul cercetării, investigării unei persoane, grup de persoane sau organizații sub aspect psihologic.

(3) *Metodele și tehnicile de asistență psihologică* constau în totalitatea demersurilor fundamentate științific, efectuate în scopul declarat al ameliorării sau optimizării activității psihice a unei persoane, grup de persoane sau organizații.

(4) *Utilizatorul de metode și tehnici de evaluare și asistență psihologică* este psihologul cu drept de liberă practică care a dobândit atestatul de liberă practică în specialitățile de competență profesională.

(5) *Testele psihologice* sunt probe sau dispozitive standardizate de extragere a unei secvențe relevante din comportamentul sau procesările cognitiv-afective ale unei persoane sau grup, fundamentate pe un etalon sau un reper pertinent și pe o interpretare responsabilă din partea psihologului cu drept de liberă practică specializat.

(6) *Producătorul* este persoana fizică sau juridică, română sau străină care este deținătoare a dreptului sau drepturilor de autor, responsabilă cu selecția, administrarea, cotarea și comercializarea metodelor și tehnicilor de evaluare și asistență psihologică.

(7) *Distribuitorul autorizat* este persoana fizică sau juridică, română sau străină, deținătoare a drepturilor de distribuție generală sau pe o anumită piață sau segment din piața metodelor și tehnicilor de evaluare și asistență psihologică.

(8) *Distribuirea pe piață* constă în acțiunea producătorului sau distribuitorului autorizat de a face disponibil contra cost sau gratuit metode sau tehnici de evaluare și asistență psihologică.

(9) *Solicitanții avizului pentru metodele și tehnicile de evaluare și asistență psihologică* sunt producătorii sau distribuitorii autorizați, care introduc spre utilizare pe piața românească metodele sau tehnicile de evaluare și asistență psihologică, altele decât cele avizate de către Comitetul director și evidențiate în *Catalogul metodelor și tehnicilor de evaluare și asistență psihologică*.

(10) *Piața* reprezintă cadrul legal reglementat prin prezentele norme precum și prin celelalte acte normative în vigoare care reglementează activitățile civile și comerciale, în care activează, pe de o parte, *producătorii și distribuitorii autorizați* de metode și tehnici de evaluare și asistență psihologică, iar pe de altă parte *utilizatorii de metode și tehnici de evaluare și asistență psihologică*.

Capitolul III

Avizarea metodelor și tehnicilor de evaluare și asistență psihologică

Art. 5 (1) În vederea obținerii *avizului pentru metodele și tehnicile de evaluare și asistență psihologică*, *solicitanții* acestuia vor depune la secretariatul *Comitetului director* un dosar pentru fiecare metodă sau tehnică de evaluare sau asistență psihologică, pentru care se solicită aviz, cuprinzând următoarele documente:

- a) cerere tip, potrivit *anexei nr. 1* la prezentele norme;
- b) actul de identitate al persoanei fizice sau actele de înregistrare ale persoanei juridice, precum și actul constitutiv al acesteia (în fotocopie, semnată și ștampilată pentru conformitate cu originalul sau după caz, tradusă și legalizată);
- c) dovada deținerii dreptului de autor sau a dreptului de distribuție asupra metodei sau tehnicii de evaluare și asistență psihologică sau orice dovadă privind întâia aducere la cunoștința publicului a metodei sau tehnicii de evaluare sau asistență psihologică, însoțită de o declarație pe proprie răspundere autenticată (original ori copie legalizată, sau după caz, tradusă și legalizată);
- d) dovada deținerii calității de *producător* sau *distribuitor autorizat*, după caz (original ori copie legalizată sau după caz, tradusă și legalizată);
- e) dovada rezervării de denumire a metodei sau tehnicii de evaluare sau asistență psihologică, după caz (original ori copie legalizată sau după caz, tradusă și legalizată);
- f) dovada efectuării cercetării de adaptare pe populația românească a metodei sau tehnicii de evaluare sau asistență psihologică (original ori copie legalizată sau după caz, tradusă și legalizată);
- g) fundamentarea teoretică a metodei sau tehnicii de evaluare și asistență psihologică (original ori copie legalizată sau după caz, tradusă și legalizată);
- h) prezentare detaliată a obiectului, scopului și destinației metodei sau tehnicii de evaluare și asistență psihologică (original ori copie legalizată sau după caz, tradusă și legalizată);

- i) proiectul privind instrucțiunile de utilizare a metodei sau tehnicii de evaluare și asistență psihologică (original ori copie legalizată sau după caz, tradusă și legalizată);
- j) lista furnizorilor români de formare profesională continuă care pot oferi pregătire teoretică și practică pentru utilizarea metodei sau tehnicii de evaluare și asistență psihologică (original, semnată și ștampilată de către solicitant);
- k) declarație de conformitate, potrivit *anexei nr. 2* la prezentele norme;
- l) studiul privind satisfacerea criteriilor de apreciere a unei metode sau tehnici de evaluare și asistență psihologică (original ori copie legalizată sau după caz, tradusă și legalizată);
- m) specialitățile de competență și treptele de specializare în care poate fi utilizată metoda sau tehnica de evaluare sau asistență psihologică, potrivit H.G. nr. 788/2005 (original sau tradusă și legalizată);
- n) bibliografia utilizată în fundamentarea teoretică (original sau fotocopie, semnată de către solicitant);
- o) prețul metodei sau tehnicii de evaluare propus și recomandat de către producător sau distribuitor, după caz (original);
- p) fotocopie după dovada achitării taxei de avizare preliminară, în cuantumul stabilit de către *Consiliul Colegiului Psihologilor din România*, denumit în continuare *Consiliul*;
- q) opis cuprinzând documentele depuse la dosar, semnat și datat de către solicitant

(2) Documentele prevăzute la alin. 1 vor fi prezentate în două exemplare, în dosare sau plicuri securizate, dintre care un exemplar în fotocopie.

(3) *Comisia Metodologică* poate solicita și alte date sau informații suplimentare.

Art. 6 În cazul solicitării avizului pentru *metodele psihologice standardizate*, fundamentarea trebuie să corespundă *Sistemului de evaluare a metodelor psihologice standardizate*, prezentat în *anexa nr. 3* la prezentele norme, elaborat de către *Comitetul director*.

Art. 7 (1) La cererea expresă a solicitantului sau reprezentantului acestuia, *Comisia metodologică* va putea stabili o dată în vederea susținerii de către solicitant a cererii de avizare.

(2) Cererea solicitantului trebuie formulată până la data atribuirii dosarului de avizare.

(3) Orice cerere ulterioară sau încercare de influențare a membrilor *Comisiei metodologice* din partea solicitanților sau reprezentanților acestora, în vederea acordării sau a neacordării avizului pentru metoda sau tehnica de evaluare sau asistență psihologică suspendă de drept procedura de avizare.

(4) Acțiunea prevăzută la alin. 3, care este întreprinsă de către un psiholog cu drept de liberă practică constituie abatere disciplinară gravă și se sancționează potrivit reglementărilor în vigoare.

Art. 8 (1) Secretariatul *Comitetului director* înregistrează cererile de avizare în ordinea cronologică a depunerii acestora.

(2) Secretariatul *Comitetului director* asigură depozitarea dosarelor de avizare până la data depunerii și înregistrării lor la *Comisia metodologică*.

(3) De la data înregistrării dosarelor la *Comisia metodologică*, responsabilitatea privind acestea aparține în totalitate membrilor comisiei.

Art. 9 (1) În termen de maxim 10 zile de la data înregistrării la secretariatul *Comitetului director* a minim 4(patru) dosare de avizare, *Comisia metodologică* va stabili data ședinței de lucru privind recepția dosarelor de avizare și atribuirea lor spre analizare membrilor comisiei.

(2) Neprezentarea membrilor *Comisiei metodologice* la lucrările acesteia constituie abatere disciplinară, cu excepția cazului fortuit sau forței majore.

(3) În cadrul ședinței de lucru, membrii comisiei fac recepția dosarelor de avizare înregistrate, precum și desemnarea membrilor comisiei care vor analiza dosarele de avizare.

(4) Verificarea inițială a dosarelor constă în verificarea datelor înscrise în opis cu documentele depuse la dosar. Orice inadvertență descoperită va fi consemnată în procesul verbal al ședinței.

Art. 10 (1) Procedura de desemnare a membrilor comisiei care vor analiza dosarele de avizare debutează cu inventarierea dosarelor și stabilirea ordinii de înregistrare.

(2) Desemnarea membrilor comisiei este efectuată în ordinea alfabetică a numelor, cu excepția președintelui comisiei care va coordona procedura de avizare.

(3) Dosarele de avizare vor fi atribuite în ordinea înregistrării lor la un număr de 2(doi) membri ai comisiei, desemnați în ordine alfabetică.

(4) Președintele comisiei încheie un proces verbal cu privire la recepția dosarelor de avizare, desemnarea membrilor comisiei, precum și a altor aspecte privitoare la procedura încheiată.

Art. 11 Aducerea la cunoștința publică sau destăinuirea de către membrii comisiei pe toată perioada desfășurării procedurilor de avizare a oricăror date cu privire la conținutul dosarelor de avizare constituie abatere disciplinară gravă și se sancționează potrivit reglementărilor legale în vigoare.

Art. 12 (1) După încheierea procedurilor de desemnare a membrilor comisiei privind analiza dosarelor de avizare a metodei sau tehnicii de evaluare și asistență psihologică, fiecărui membru i se înmânează un exemplar din dosarul de avizare repartizat spre analizare, fapt ce se consemnează în procesul verbal al ședinței.

(2) Membrii comisiei stabilesc prin vot data următoarei ședințe de lucru, dar nu mai târziu de *20 de zile calendaristice*, fapt ce se consemnează în procesul verbal al ședinței de lucru.

(3) La data stabilită pentru următoarea ședință de lucru, fiecare membru al comisiei va elabora și prezenta un raport privind avizarea sau neavizarea metodei sau tehnicii de evaluare sau asistență psihologică analizată.

(4) După audierea fiecărui raport, membrii comisiei votează în plen propunerea de avizare sau neavizare a metodei sau tehnicii de evaluare sau asistență psihologică.

(5) Hotărârea membrilor comisiei este luată cu majoritatea voturilor valide. Voturile de abținere nu sunt considerate valide. Fiecare membru al comisiei este obligat să exprime un vot valid.

Art. 13 (1) În cazul în care membrii comisiei nu consideră concludente documentele prezentate, aceștia vor comunica solicitanților o adresă prin care se precizează probele, documentele sau alte clarificări pe care aceștia trebuie să le prezinte în completare.

(2) Termenul de completare a dosarelor de avizare este stipulat în adresa comunicată solicitantului. Data comunicării este considerată data poștei.

(3) Datele privind completarea dosarului vor fi analizate după procedura descrisă în prima ședință a *Comisiei metodologice*.

(4) Neprezentarea în termen a documentelor de completare este echivalentă cu renunțarea la cererea de avizare.

Art 14 (1) *Comisia metodologică* propune *Comitetului director* acordarea avizului pentru metoda sau tehnica de evaluare sau asistență psihologică analizată sau respingerea cererii de avizare. În baza propunerii înaintate de către *Comisia metodologică*, *Comitetul director* hotărăște eliberarea *avizului*, conform *anexei nr. 4* la prezentele norme sau respingerea cererii de avizare.

(2) Termenul de eliberare a avizului pentru metoda sau tehnica de evaluare sau asistență psihologică este de *30 de zile* de la data înregistrării propunerii de avizare la secretariatul *Comitetului director*.

(3) Adresa de respingere a cererii de avizare a metodei sau tehnicii de evaluare sau asistență psihologică va fi comunicată solicitantului în termen de *30 zile* de la data înregistrării propunerii de respingere a cererii la secretariatul *Comitetului director*.

(4) Solicitantul de aviz poate contesta respingerea cererii de acordare a avizului în termen de *15 zile* de la data comunicării adresei de respingere a cererii.

(5) Hotărârea dată în soluționarea contestației este definitivă. De asemenea, hotărârea *Comitetului director*, necontestată în termen este considerată acceptată de către solicitanți.

Capitolul IV

Tipurile de avize

Art. 15 (1) În urma analizei efectuate de către membrii comisiei, *Comisia metodologică* poate propune *Comitetului director* acordarea a *două tipuri de avize*:

- a) *aviz pe perioadă nedeterminată*;
- b) *aviz temporar*, până la încheierea anului calendaristic următor.

(2) *Comisia metodologică* va întocmi o motivare prin care să justifice tipul de aviz dat și recomandările făcute pentru îmbunătățirea calităților psihometrice ale probelor. De asemenea, vor fi aprobate și eliberate *instrucțiunile de utilizare* ale metodei sau tehnicii de evaluare sau asistență psihologică.

(3) *Avizul eliberat pe perioadă nedeterminată* corespunde unei metode sau tehnici de evaluare sau asistență psihologică care îndeplinește toate condițiile de calitate necesare utilizării de către psihologii cu drept de liberă practică. Metoda sau tehnica de evaluare sau asistență psihologică avizată pe perioada nedeterminată este supusă reavizării.

(4) *Avizul temporar* corespunde unei metode sau tehnici de evaluare sau asistență psihologică care îndeplinește condițiile minime de calitate necesare utilizării de către psihologii cu drept de liberă practică. *Avizul temporar* nu este supus reavizării, fiind valabil până la data de 31 decembrie a anului calendaristic următor.

(5) Respingerea cererii de acordare a avizului pentru metoda sau tehnica de evaluare sau asistență psihologică corespunde neîndeplinirii condițiilor minime de calitate necesare utilizării de către psihologii cu drept de liberă practică sau prezentării unor date incomplete.

Art. 16 (1) Avizul eliberat de către *Comitetul director* se rezumă la aria de aplicabilitate specificată în caracteristicile domeniului vizat, și nu poate fi extins spre alte domenii de activitate fără o avizare prealabilă.

(2) Sunt exceptate de la necesitatea obținerii avizului metodele sau tehnicile de evaluare sau asistență psihologică aflate în perioadă de experimentare, ale căror rezultate nu pot fi utilizate în procesul de testare psihologică.

(3) Dovada avizării metodei sau tehnicii de evaluare și asistență psihologică o constituie hotărârea *Comitetului director*.

Art. 17 (1) Psihologii cu drept de liberă practică utilizează în desfășurarea activității numai metodele sau tehnicile de evaluare sau asistență psihologică avizate în condițiile prezentelor norme.

(2) Încălcarea dispozițiilor alin. 1 reprezintă abatere disciplinară gravă și este sancționată potrivit actelor normative în vigoare.

(3) Prezentarea de către orice persoană a unor metode sau tehnici de evaluare sau asistență psihologică neavizate de către *Comitetul director* ca fiind avizate de către acesta se sancționează potrivit legii penale.

Capitolul V

Dispoziții finale

Art. 18 Prezentele *norme* reprezintă procedura unică de avizare a metodelor sau tehnicilor de evaluare sau asistență psihologică.

Art. 19 Prezentele *norme* au fost adoptate de către *Comitetul director* în ședința din data de 8 iunie 2006, cu respectarea dispozițiilor legale în vigoare.

Anexa nr. 1

Domnule Președinte,

Subscrisa/Subsemnata/-ul _____, cu domiciliul în _____, legitimat/-ă cu ____, seria ____, nr. _____, CNP: _____, în calitate de _____ (producător/distribuitor autorizat), reprezentat/-ă prin _____, în calitate de _____, vă rog să avizați metoda sau tehnica de evaluare sau asistență psihologică, cu denumirea : _____, pentru a putea fi utilizat în specialitatea: _____, de către psihologii cu drept de liberă practică.

În acest scop, potrivit art. 5 din *Normele de avizare a metodelor și tehnicilor de evaluare și asistență psihologică*, anexează la dosar documentele necesare pentru avizare.

Data: _____

Semnatura: _____

Anexa nr. 2

Emitent: _____
Sediul social: _____
Cod de înregistrare fiscală/C.U.I.: _____
Tel./Fax _____

DECLARAȚIE DE CONFORMITATE

Noi, _____,
(denumirea completă a persoanei juridice sau persoanei fizice autorizate)

având certificat de înregistrare/autorizație nr. _____/_____, asiguram, garantam și declarăm pe propria răspundere, conform prevederilor *art. 5 din Hotărârea Guvernului nr. 1.022/2002* privind regimul produselor și serviciilor care pot pune în pericol viața, sănătatea, securitatea muncii și protecția mediului, că produsul/serviciul

_____ (denumirea, tipul sau modelul, eventual numărul lotului, seriei și numărul de exemplare)

la care se referă prezenta declarație nu pune în pericol viața, sănătatea, securitatea muncii, nu produce un impact negativ asupra mediului și este în conformitate cu: _____

_____ (titlul și/sau numărul și data publicării documentului/documentelor normativ/normative)

_____ (locul și data emiterii)

_____ (numele și prenumele în clar și ștampila)

Anexa nr. 3

SISTEMUL DE EVALUARE A METODELOR PSIHOLOGICE STANDARDIZATE

GHID PENTRU SOLICITANȚI

Sistemul de evaluare a calității unui test cuprinde *cinci criterii*:

- a) **Baza teoretică și soliditatea procedurilor de dezvoltare / perfecționare a testului;**
- b) **Calitatea materialelor pentru testare și claritatea manualului;**
- c) **Reguli / etaloane;**
- d) **Fidelitatea;**
- e) **Validitatea;**

Pentru fiecare criteriu a fost întocmit câte *un chek-list* cu un număr de itemi. Unii itemi (cel puțin câte unul pentru fiecare criteriu) sunt așa numitele întrebări cheie, care verifică dacă anumite condiții minimale sunt îndeplinite. Dacă o întrebare cheie este evaluată negativ, evaluarea pentru acel criteriu automat va fi “insuficient”. Pentru a spori o interpretare uniformă a itemilor de către evaluatori și pentru a explica sistemul utilizatorilor de teste și celor care dezvoltă testele, secțiunile cu comentarii furnizează informații detaliate asupra evaluării și aprecierii itemilor. O dată ce itemii au fost evaluați, semnificațiile finale (“**insuficient**”, “**suficient**”, “**bun**”) pentru cele cinci criterii sunt stabilite prin semnificația regulilor de apreciere.

În termenii normelor de avizare a testelor psihologice, dacă la unul din cele cinci criterii proba dobândește calificativul « insuficient » nu va fi avizată de către CPR și nu va fi recomandată spre utilizare psihologilor cu drept de liberă practică.

Dacă o probă primește la fiecare criteriu evaluat cel puțin calificativul suficient ea va fi avizată de către Comisia Metodologică și va fi recomandată spre uz, psihologilor cu drept de liberă practică, după cum urmează :

- proba primește aviz definitiv dacă dintre cele cinci criterii evaluate, cel puțin trei sunt evaluate ca fiind « bun », celelalte primind calificativul suficient. Între criteriile care obțin calificativul bun trebuie neapărat să se regăsească criteriul « validității » probei respective.
- în toate celelalte situații neincluse mai sus, în condițiile în care o probă va primi cel puțin calificativul suficient la fiecare criteriu evaluat, se va acorda un aviz temporar probelor respective.

Acest document este constituit din două secțiuni. În prima secțiune sunt prezentate **întrebările pentru fiecare din cele cinci criterii**. În a doua parte sunt descrise **explicațiile generale pentru fiecare criteriu**, explicațiile cu referire la fiecare item și regulile de combinare a itemilor evaluativi (pentru a se obține evaluarea pentru fiecare criteriu).

CRITERII ȘI LISTA DE ÎNTREBĂRI**1. Bazele teoretice ale testului**

- 1.1 Întrebare cheie: Sunt specificate scopurile și domeniile de aplicare ale testului? (dacă evaluarea acestui item este negativă, se trece direct la itemul 2.1)
- 1.2 Este descris motivul care a stat la baza construirii testului și conceptele pe care testul le măsoară sunt clar definite?
- 1.3 Relevanța conținutului testului este justificată față de conceptele măsurate?

2. Calitatea materialelor de testare și a manualului testului**2A. *Materialele testului***

- 2.1. Întrebare cheie: Sunt standardizați itemii?
- 2.2.a. Întrebare cheie: Există un sistem obiectiv de scorare? sau
- 2.2.b. Există specificat un sistem clar și complet pentru evaluare sau examinare în cazul în care testul trebuie să fie scorat de către evaluatori sau examinatori?
- 2.3. Întrebare cheie: Itemii testului au un conținut fără specificații rasiste sau alte mențiuni ofensive pentru anumite categorii de oameni? (Dacă evaluarea unuia din itemii prezentați mai sus este negativă se trece la itemul 2.7)
- 2.4.a. Itemii, broșura testului, scalele de răspuns și foile de răspuns sunt în așa fel încât erorile de completare să fie evitate?
- 2.4.b. Care este calitatea materialelor folosite pentru testare?
- 2.5. Sistemul de scorare este gândit în așa fel încât erorile să fie evitate?
- 2.6. Instrucțiunile pentru cel care primește testul sunt complete și clare?

2B. *Manualul testului*

- 2.7. Întrebare cheie: Este disponibil manualul testului? (dacă evaluarea la acest item este negativă se trece la itemul 3.1)
- 2.8. Instrucțiunile pentru cel care administrează testul sunt complete și clare?
- 2.9. Este specificat în ce gen de situații poate fi aplicat testul și care sunt limitele testului?
- 2.10. Există publicat în manual un rezumat al cercetării?
- 2.11. Este ilustrat (de exemplu, prin menționarea câtorva cazuri) modul în care scorurile testului ar putea fi interpretate?
- 2.12. Este indicat ce tip de informații sunt importante pentru interpretarea scorurilor testului?
- 2.13. Este specificat ce calificări profesionale sunt cerute pentru a administra și interpreta testul?

3. Etaloanele

- 3.1. Întrebare cheie: Etaloanele sunt prevăzute (incluzând tabelele și scorurile fixate)
- 3.2. Întrebare cheie: Care este calitatea datelor standard și a studiului de etalonare?
- 3.3. Este indicat pentru care alte categorii de grupuri, etaloanele pot fi generalizate, și dacă da, sunt menționate marjele de eroare?
- 3.4. Sunt explicate semnificația și limitele tipului de scală sau scorurilor care sunt folosite și scala se dovedește a fi compatibilă cu scopul stabilit al testului?
- 3.5. Există informații despre medii, abateri standard și distribuția scorurilor?
- 3.6. Există informații despre erorile standard de măsurare și/sau erorile standard de estimare a intervalului de încredere adecvat al măsurării?

- 3.7. Există informații despre posibilele diferențe dintre subgrupuri (de exemplu, referitoare la sexul subiecților sau apartenența etnică)?
- 3.8. Este raportat anul în care au fost colectate datele pentru fiecare grup în vederea etalonării?

4. Fidelitatea

- 4.1. Întrebare cheie: Există informații referitoare la fidelitatea testului? (dacă evaluarea itemului este negativă se trece la itemul 5.1)
- 4.2. Rezultatele din cadrul studierii fidelității sunt suficiente raportat la tipul deciziilor pe care se bazează testul?
 - a. Fidelitatea formelor paralele
 - b. Consistența internă
 - c. Fidelitatea test-retest
 - d. Fidelitatea inter-evaluatori
- 4.3. Care este calitatea investigării fidelității?
 - a. Sunt corecte procedurile de calculare a coeficienților fidelității?
 - b. Sunt compatibile eșantioanele pentru calcularea coeficienților fidelității cu scopul stabilit al testului?
 - c. Este posibil să se realizeze o judecată completă a fidelității testului pe baza informațiilor date?

5. Validitatea

5.A. *Validitatea de construct*

- 5.1. Întrebare cheie: Există informații referitoare la validitatea de construct a testului? (Dacă evaluarea la acest item este negativă se trece la itemul 5.4).
- 5.2. Rezultatele privind studierea validității susțin semnificația propusă a conceptelor (sau: Rezultatele privind studierea validității arată clar ceea ce este măsurat)?
- 5.3. Care este calitatea studiilor de validare a constructului (constructelor) măsurate?
 - a. Procedurile folosite pentru obținerea și calcularea datelor privind validitatea de construct sunt corecte?
 - b. Eșantioanele folosite în studierea validității de construct sunt compatibile cu scopul stabilit al testului?
 - c. Care este calitatea altor instrumente folosite în studierea validității de construct?
 - d. Este posibilă luarea unei hotărâri corecte privind validitatea de construct a testului pe baza informațiilor date?

5.B. *Validitatea relativă la criteriu*

- 5.4. Întrebare cheie: Există informații despre relația test – criteriu? (Dacă evaluarea la acest item este negativă, itemii 5.5 și 5.6 pot fi săriți).
- 5.5. Rezultatele din cadrul studierii validității sunt suficiente în ceea ce privește tipul deciziilor pentru care a fost propus testul?
- 5.6. Care este calitatea studiului validității de criteriu?
 - a. Procedurile folosite pentru obținerea și calcularea datelor privind criteriul validității sunt corecte?

- b. Eșantioanele folosite în studierea validității de criteriu sunt compatibile cu scopul stabilit al testului?
- c. Care este calitatea studiului validității de criteriu (predictivă)?
- d. Este posibilă luarea unei hotărâri corecte privind validitatea de criteriu a testului pe baza informațiilor date?

EXPLICAȚII ȘI REGULI DE APRECIERE

1. Bazele teoretice ale testului

Explicații generale

Construcția testului solicită o pregătire amănunțită. În cele din urmă, scorurile testului constituie temeiul în care se elaborează judecăți referitoare la diferențele intra-individuale (în cadrul unor contexte precum sistemul educațional și consilierea vocațională), diferențele inter-individuale (de exemplu, în selecția personalului), și diferențele între grupuri sau situații (precum cele din diagnozele organizaționale). Informațiile asigurate de autorul testului ar trebui să permită viitorului utilizator al testului să discearnă dacă testul este potrivit pentru scopul pe care îl urmărește. Pornind de la aceasta, primul pas constă în a furniza o descriere clară a conceptului teoretic pe care testul își propune să-l măsoare. Alegerea conținutului testului și metodele prin care un concept este măsurat ar trebui de asemenea, luate în considerare. Aceste aspecte depind doar de viabilitatea teoriei și minuțiozitatea descrierii ipotezelor de bază. Problema cu referire la elaborarea cu succes a testului ca o construcție demnă de încredere și validă sunt luate în considerare în cadrul criteriilor 3, 4, 5 focalizate pe calitatea și rezultatele cercetării.

Explicații / recomandări (explicații) pentru fiecare item

Itemul 1.1 Construcția testului ar trebui să înceapă cu o reflecție asupra scopului testului. Scopul este de a prezice anumite comportamente (predictiv) ? Are intenția de a evalua progresul educațional sau efectele unui training (monitorizare)? Este utilizat pentru a aprecia potrivirea pentru un anumit tip de tratament (diagnostic) etc.?

Este de asemenea esențial să se menționeze grupul sau grupurile pentru care este realizat testul, detaliind aspecte precum vârsta grupului, ocupația, nivelul studiilor și dacă este vorba de un context luat în considerare este normal sau clinic. Cu cât sunt mai ambițioase scopurile autorului testului, cu atât mai mari sunt obligațiile sale de a arăta date empirice precum etaloane și dovada validității instrumentului.

Itemul 1.2 Autorul testului ar trebui să clarifice dacă testul reflectă o teorie existentă sau dacă acesta este bazat pe o dezvoltare nouă care poate aduce eventual, schimbări minore sau majore în cadrul altei teorii. Indiferent de teoria pe care se bazează testul, aceasta trebuie descrisă detaliat. Dacă testul este o traducere sau adaptare a unui instrument străin, trebuie asigurată o informație de fond suficientă, o listă de referințe nefiind suficientă. Chiar și atunci când testul își propune să măsoare constructe foarte cunoscute precum inteligența, de exemplu, o definiție a conceptului trebuie dată pentru a se lămurii ce aspecte ale comportamentului aparțin domeniului respectiv. De asemenea, trebuie descrise și anumite asemănări sau diferențe cu teste similare. Din această prezentare, valoarea adăugată a noului instrument ar trebui să fie evidentă.

Itemul 1.3 Această întrebare se referă la pașii prin care conceptul ce se măsoară este operaționalizat. Definierea ariei de conținut trebuie să facă posibil a se stabili care itemi aparțin domeniului și care nu. Acest aspect poate fi realizat prin analizarea conceptului într-un mod care să clarifice care din fațete este luată în considerare. Considerațiile teoretice sau conținuturile asociate pentru greutatea acordată acestor fațete ar trebui și ele consemnate, și problema itemilor luați în considerare pentru eşantionare ar trebui menționată. Chiar dacă

itemii sunt eliminați sau modificați în cursul construirii sau adaptării testului, consecințele acestor schimbări pentru măsurarea conceptului original trebuie indicat (domeniul conținut poate fi schimbat, îngustat sau incomplet acoperit ca rezultat al acestor modificări).

2. Calitatea materialelor pentru testare și a manualului testului

Explicații generale

Scorul la un test poate fi interpretat ca o măsură de încredere dacă testul este administrat în condiții standardizate. Scopul standardizării situației de testare este să se preîntâmpine apariția unor factori necontrolați care să afecteze scorul. De exemplu, administrarea și instruirea ar trebui să fie standardizate într-o asemenea măsură încât influența administratorului de test și a variațiilor în instrucțiuni să fie eliminată sau minimizată.

Explicații / recomandări pentru fiecare item (Criteriul 2A)

Itemul 2.1 Itemii testului sunt standardizați atunci când ei sunt aceeași pentru fiecare respondent din punct de vedere al conținutului, formei și succesiunii. Standardizarea este o condiție importantă pentru interpretarea și compararea scorurilor. O excepție, care ține de cerința unei succesiuni uniforme a itemilor testului este acceptată în cazul testelor adaptive. Totuși, și pentru acest tip de test, regulile de stabilire a selecției oricărui dintre itemi trebuie redate explicit.

Itemul 2.2.a Un sistem de scorare este considerat obiectiv atunci când valoarea scorului acordată pentru toate posibilele răspunsuri la toți itemii testului sunt stabilite în așa mod, încât orice persoană calificată care scorează itemii, cu excepția erorilor de redactare, acordă exact același scor pentru același răspuns. Acest aspect este aplicabil în mod particular în cazul testelor de abilitate, creion – hârtie și chestionarelor cu itemi cu alegere multiplă.

Itemul 2.2.b Pentru scalele de observare, testele proiective, subtestele bateriei individuale de inteligență și întrebările cu răspunsuri deschise, cotașul nu poate fi strict obiectiv. Pot fi descrise însă dovezile care asigură un nivel cât mai ridicat de obiectivitate în cotaș.

Itemul 2.3 Conținutul unui test este calificat drept rasist dacă conține cuvinte, fraze, imagini sau alte elemente folosite în itemi sau instrucțiuni aduceau ofense membrilor unui sau mai multor grupuri etnice minoritare. Conținutul unui test este calificat drept ca discriminatoriu din punct de vedere al grupurilor etnice, atunci când testul conține cuvinte, idiomuri sau imagini complicate cu specific cultural. Principiul utilizării restricționate poate fi aplicat la fel de bine și altor grupuri. Un exemplu, este un inventar de interese care conține imagini cu itemi-stereotipi legați de sexul personajelor. Acest principiu nu se aplică testelor care sunt destinate să măsoare concepte raportate la rasism sau sex (așa cum este Scala F a lui Adorno sau scala pentru androgenie).

Itemul 2.4.a Concluziile care trebuie avute în vedere în evaluarea acestui item includ: (a) claritatea itemilor sau distribuțiile pentru grupurile pentru care testul este stabilit (majoritatea itemilor nu trebuie să treacă dincolo de restricțiile acestor grupuri) și (b) foile de

răspuns, dacă se folosesc, ar trebui realizate într-un mod în care erorile de completare (cum ar fi sărirea peste un item) pot fi detectate rapid.

Itemul 2.4.b Acest item vizează toate aspectele practice ale materialelor testului care nu au fost acoperite în totalitate de itemii de la criteriul 2A. Acestea include următoarele considerații:

- Itemii sunt formulați într-un limbaj obișnuit?
- Testul este clar?
- Pot fi deosebite clar culorile sau simbolurile (dacă se folosesc) unele față de altele (chiar și pentru indivizii daltoniști)?
- Materialele testului sunt durabile?

Itemul 2.5 În evaluarea acestui item, atenția trebuie acordată următoarelor aspecte:

- Procedura de cotare trebuie descrisă în totalitate și clar
- Dacă se folosesc grile pentru calcularea scorului, trebuie să existe instrucțiuni clare asupra modului în care se plasează aceste grile deasupra foilor de răspuns.
- Dacă se folosesc grile pentru calcularea scorului, acestea trebuie să se potrivească cu acuratețe peste foile de răspuns.
- Dacă se folosesc grile pentru calcularea scorului, versiunea testului trebuie imprimată pe grilele de corectare. Acesta este, în mod deosebit, important atunci când testul este revizuit.
- Ar trebui să existe instrucțiuni clare pentru scorarea itemilor pierduți (săriți).
- Trebuie dată o indicație asupra a cât de mulți itemi pot fi pierduți fără ca testul să își piardă valoarea.
- Dacă sunt implicați evaluatori sau observatori, ar trebui să existe instrucțiuni clare despre cum să se facă față diferențelor dintre evaluatori și observatori,

În general, folosirea de foi de răspuns separate este preferabilă, deoarece în urma scorării unor foi de răspuns vor rezulta mai puține erori decât scoarea unei broșuri de test.

Itemul 2.6 În acest sistem de evaluare este făcută o distincție între instrucțiunile pentru cel căruia i se administrează testul și pentru cel care administrează testul. Calitatea instrucțiunilor pentru cel testat sunt evaluate în acest item; instrucțiunile pentru administratorul testului sunt evaluate în itemul 2.8. Instrucțiunile sau explicațiile pentru cel testat se găsesc în materialele testului și de obicei constituie prima sau primele pagini ale broșurii testului sau prima / primele pagini de text atunci când testul este administrat pe computer. Instrucțiunile trebuie să fie standardizate și ar trebui să includă următoarele elemente:

- unul sau mai multe exemple de întrebări
- informații despre cum se înscriu (sau se redactează, în cazul testării computerizate) răspunsurile.
- O strategie de a presupune sau de a răspunde, când alternativele au același grad de probabilitate sau aplicabilitate.
- Limite de timp

Explicații / recomandări per item (Criteriul 2B)

Itemul 2.7 Autorul unui test trebuie să furnizeze un manual care să conțină informații practice (numit în mod obișnuit “Ghidul utilizatorului”) și informații tehnice (numit în mod obișnuit Manualul tehnic). Dezbaterile, articolele de revistă sau lucrările de cercetare nu ar trebui privite ca un manual.

Itemul 2.8 Obiectivul principal al recomandărilor pentru administratorul de test în manual este să asigure standardizarea testului. Principiile ar trebui să fie, pe cât de posibil, explicite în descrierea lor, în privința a ceea ce are de spus exact administratorul, ce nu trebuie să facă sau să spună și sarcinile pe care acesta trebuie să le îndeplinească (cum ar fi de exemplu, aranjarea materialelor într-o anumită ordine pentru un test de abilitate). O afirmație de genul “administratorul testului explică scopul testului celui căruia îi este administrat” este considerată insuficientă. Manualul testului trebuie să furnizeze de asemenea, un plan detaliat asupra modului cum să se descurce cu întrebările obișnuite pe care le pot adresa subiecții testați.

Itemul 2.9 Un manual trebuie să fie complet, precis și clar în privința aplicabilității testului. Acesta poate ghida către sugestii de utilizare diferită a testului, în funcție de situații sau aplicații specifice. Exemple în acest sens ar fi următoarele:

- S-a specificat faptul că deciziile legate de clasificarea educaționale nu ar trebui luate doar pe baza unui singur scor de test?
- A fost menționată relația dintre scorul testului și ulterioarele procese de învățare în cazul evaluării progresului?
- Rezultatele obținute la un test într-o situație clinică pot conduce la concluzii întemeiate empiric sau ele servesc doar ca ipoteze de cercetare?
- S-a evidențiat faptul că scorurile la un singur test nu ar trebui folosite ca un temel în luarea unor decizii privind consilierea vocațională?
- În cazul testelor pentru selecție de personal, au fost indicate categoriile ocupaționale pentru care testul a fost destinat și specificate capacitățile semnificative pentru aceste ocupații?

Itemul 2.10 Pentru utilizatorii de teste, manualul va constitui principala sursă de informații, deoarece mulți dintre aceștia nu vor avea acces a dezbateri, raporturi de cercetare sau alte materiale publicate și nu vor avea timp (sau nu vor fi suficient de interesați) să citească toate detaliile tehnice. Un sumar al etalonării, studiilor pentru fidelitate și validitate trebuie așadar prevăzute în manual. Dacă noi cercetări furnizează informații suplimentare folositoare, utilizatorii ar trebui informați prin mijloace auxiliare sau revizuri ale manualului. În acest item doar prezența informației în manual este solicitată. Calitatea design-ului de cercetare și rezultatele sunt evaluate în Criteriile 3, 4 și 5. Dacă acest tip de informații nu este publicat în manual, acest item va primi o evaluare negativă.

Itemul 2.11 În general, includerea unui număr de cazuri descriptive în manual ar trebui să ajute utilizatorul testului în interpretarea scorurilor obținute.

Itemul 2.12 A fost dată vreo indicație dacă există alte variabile care contribuie la predicție?

Itemul 2.13 Mențiunile privind competențele utilizatorilor ar trebui să detalieze pregătirea specifică, certificarea profesională sau experiența necesară.

3. Etaloanele

Explicații generale

Luat ca atare, rezultatul brut este puțin semnificativ, sau chiar nesemnificativ. În general, scorul brut poate fi înțeles prin raportarea la un etalon. Etalonul poate deriva dintr-o arie de aptitudini sau aspectele stăpânite de subiecți (la care se referă conținutul interpretării) sau poate fi bazat pe distribuția scorurilor prin raportare la un grup (interpretare prin raportarea la un etalon de referință). În interpretarea raportată la o tablă de corecție, rezultatul fiecărui subiect testat este examinat separat, pentru a vedea care răspunsuri sunt corecte și care sunt greșite. Rezultatele nu sunt comparate cu cele ale altui subiect. Pe de altă parte, în interpretarea bazată pe etalon, compararea cu ceilalți este principiul de bază al acestui proces. Aici, scorul subiectului testat este comparat cu scorurile altor indivizi care pot fi folosite în comparare (adică, pe baza unor similarități de vârstă, studii, profesie, etc). Un etalon, fie într-un caz, fie în celălalt constituie o cerință de bază a tuturor testelor. Însă, pot exista excepții, mai ales în cazurile testelor în care numai interpretarea intra-individuală este recomandată. În asemenea situații, întrebarea la acest capitol, nu necesită răspuns și calificativul “nu este aplicabilă” poate fi utilizat.

Etaloanele sunt susceptibile la degradare. Dintre toate proprietățile psihometrice ale unui test, etaloanele sunt cele mai sensibile la factori precum schimbări sociale, educaționale sau profesionale. Prin urmare, fie un test trebuie să fie reetalonat din când în când, fie autorul testului trebuie să arate prin mijloace de cercetare că nu este necesară o reetalonare a testului. Pentru a atrage atenția utilizatorului de teste asupra posibilității că etaloanele pot fi depășite, mențiunea “Etaloanele sunt depășite.” este adăugată la evaluarea testului, dacă datele din etalon sunt completate cu mai mult de 15 ani în urmă. După trecerea altor 5 ani în care nu a fost realizată reetalonarea, mențiunea anterioară se va schimba în “Etaloanele nu mai sunt utilizabile, deoarece sunt depășite.” Aceste specificații sunt completate o dată pe an.

Explicații / recomandări pentru fiecare item

Itemul 3.1 Etaloanele trebuie să fie valabile la momentul în care testul este publicat pentru uzul actual. Datele de etalonare bazate pe scoruri individuale nu sunt relevante pentru testele proiectate pentru etaloane raportate la evaluările de grup și viceversa. Etaloanele nu mai pot fi folosite atunci când au fost făcute modificări ale testului, de exemplu în cazul modificărilor de itemi sau instrucțiuni. Transformarea unui test creion-hârtie într-un test computerizat este de obicei puțin influențată din această perspectivă în cazul scorurilor la chestionare de personalitate, dar în cazul testelor de performanță, de obicei, noi date de etalonare trebuie adunate. Aceasta mai ales, în cazul în care sunt implicate limite de timp.

Itemul 3.2 Etaloanele de bază trebuie să fie prezentate pentru toate scopurile și aplicațiile pentru care testul este recomandat de către autor (vezi itemul 1.1). Poate reieși faptul că grupurile pentru care etaloanele sunt prezentate să acopere doar o parte din aplicațiile proiectate. De exemplu, când autorul unui test indică faptul că un test este stabilit atât pentru consilierea vocațională spre școlile tehnice, cât și pentru selecția pentru meserii

tehnice, etalonul ar trebui să fie furnizat pentru ambele situații. Totuși, nu ar fi realist să se solicite etaloane pentru fiecare profesiune tehnică. Un etalon de grup trebuie să îndeplinească anumite cerințe pentru a atinge scopul său (acesta este de a furniza un set de puncte de referință de încredere). În primul rând trebuie să fie reprezentativ pentru grupul de referință și în al doilea rând eșantionul trebuie să fie suficient de mare. Pentru a facilita o evaluare a reprezentativității eșantionului pentru etalonare, trebuie date o definiție a populației și o descriere completă a structurii eșantionului și a procesului de colectare a datelor. Trebuie să fie evident din descrierea populației avute în vedere, dacă datele sunt colectate local sau național, dacă strângerea datelor acoperă media populației sau numai a anumitor indivizi cu calități specifice (de exemplu persoane cu probleme mentale sau cu o anumită pregătire școlară de bază), ce reguli de eșantionare au existat, ș.a.m.d.

Cel mai adesea, datele sunt colectate prin așa numitele “eșantioane la îndemână”. Asemenea eșantion poate consta în clienții care apelează la un serviciu de consiliere vocațională, studenți la psihologie sau pacienți din centre medicale de psihiatrie, de exemplu. În general, eșantioanele de acest tip oferă un etalon sărac deoarece motivul testării este raportat la scorul testului. De aceea, aceste eșantioane nu pot fi considerate reprezentative pentru populația țintă (care, în exemplele date mai sus, se referă doar la liceeni, studenți, respectiv persoane cu probleme psihiatrice).

Itemul 3.3 Colectarea datelor pentru etalonare este o procedură costisitoare și laborioasă, care solicită să nu fie realizată pentru toate grupurile posibile. Etaloanele pentru grupurile cu vârste medii pot fi obținute prin mijloace de extrapolare sau etaloane care pot fi generalizate pentru profesii sau ocupații similare. Autorul testului trebuie să justifice extrapolarea sau generalizarea etaloanelor prin descrierea punctelor critice ale asemănării dintre grupuri (de exemplu, în termeni de conținut al profesiei sau al educației) sau prin furnizarea de date din cercetare (cum ar fi scorurile la variabile înrudite).

Când vârsta sau clasele etaloanelor sunt furnizate, amplitudinea vârstei și nivelul intervalului contribuie la erori (bias-uri) în scorurile testului: performanța este subestimată la începutul intervalului și supraestimată la sfârșitul acestuia. Testele de abilitate pentru copiii mici sunt, în mod particular, mai vulnerabile la acest efect, cu o variație de 10 sau mai multe puncte QI în cadrul unei perioade de un an. Chiar în primii ani ai școlii generale, diferențele dintre doi ani consecutivi pot depăși tot atât de mult, cât jumătate din deviația standard. Acest tip de bias poate fi împiedicat prin extinderea listelor de etaloane. Dacă testul este proiectat pentru a fi folosit într-o anumită perioadă a anului, aceasta trebuie menționat clar și datele pentru etalonare trebuie adunate tot în aceeași perioadă. În orice caz, vârsta medie în luni sau perioada anului în care etaloanele sunt colectate trebuie să fie menționate pentru vârsta și educația la care se raportează etaloanele.

Itemul 3.4 Pentru transformarea scorurilor brute în scoruri standard există un sistem numeros de categorii de scale, din care se poate alege. O alegere poate fi făcută între scorurile standard, percentile și o scală cu mai multe sau mai puține unități. Pentru o situație particulară, autorul testului poate decide să proiecteze un nou sistem sau să adapteze unul existent. Când scopul testului trece sub categoria “important”¹ este indicat să se aleagă un sistem precis cu mai multe unități. Totuși, pentru testele din această categorie este recomandată folosirea intervalului de încredere. A opta pentru un sistem brut, înseamnă a sacrifica precizia, dar pot face rezultatele mai ușor și rapid de înțeles (ex. apt vs. inapt). Indiferent de sistemul de scală utilizat, ar trebui să fie descrise și motivele pentru alegerea scalei ar trebui să fie menționate.

¹ Decizii importante sunt: decizii luate pe baza scorului unui test care sunt esențiale sau în termen scurt, ireversibile și asupra cărora cel testat are o mică influență.

Itemul 3.5 Aceste date trebuie furnizate pentru fiecare etalon de grup. Aspecte precum boltirea, simetria sau bimodalitatea distribuției scorurilor sunt relevante, precum și diferențele posibile între acești parametri între etaloanele de grup. De exemplu, poate fi cazul în care scorurile la un chestionar sunt mai mult sau mai puțin distribuite normal într-un grup, în timp ce 50% din participanții dintr-un alt grup obțin scoruri scăzute. Alte informații utile ar putea fi existența unor efecte de bază sau de plafon („floor and ceiling effects”) obținute într-un test de abilități cognitive. Un utilizator de test are nevoie de astfel de informații pentru a interpreta corect scorurile la test.

Itemul 3.6 În literatura de specialitate, eroarea standard de măsurare și eroarea standard estimată nu sunt diferențiate. Eroarea standard de măsurare dă un interval de încredere care este simetric în jurul scorurilor observate. Această indică faptul că scorurile observate diferă semnificativ de fiecare scor adevărat care se află în acest interval. Formula pentru calcularea erorii standard de măsurare este: $\sigma_{est} = \sigma_x \sqrt{1-r_{xx}}$. Intervalul de încredere este important atunci când este cerută utilizarea semnificației testelor (scorul unei persoane A diferă de scorul persoanei B sau față de scorul medie X_0 ?). Pe baza acestor erori standard se pot calcula intervalele de încredere ce cuprind cu o probabilitate foarte mare scorul adevărat al subiectului testat. Pentru testele încadrate în categoria “importante”, intervalele de probabilitate trebuie să fie prevăzute, pentru a obține o evaluare pozitivă a acestui item. Dacă este dată doar eroarea standard va fi dată o evaluare ‘+/-’. Pentru teste mai puțin importante, raportarea erorii standard corecte a măsurării este suficientă pentru a se acorda o evaluare pozitivă.

Itemul 3.7 Există diferite motive pentru care diferențele dintre subgrupuri trebuie studiate și raportate:

- rezultatele pot indica un impact opus;
- rezultatele pot justifica bias-urile de cercetare raportat la test sau la item;
- utilizatorul testului primește date care îi permit să decidă dacă să ia sau nu în considerare posibilele diferențe;

Aceste studii se aplică doar acolo unde subgrupurile sunt relevante pentru scopul testului, așa cum sunt grupurile împărțite pe sex, vârstă sau apartenența etnică.

Itemul 3.8 Menționarea anului în care au fost colectate datele pentru etalonare este importantă pentru a decide dacă etaloanele sunt depășite, și dacă da, cu cât timp.

4. Fidelitatea

Explicații generale

Varianța în scorurile unui test constă în varianță adevărată și varianță eroare. Sursele varianței eroare pot fi diferite. Numeroșii coeficienți de fidelitate care pot fi calculați reflectă aceste surse în diferite grade. De aceea, nu este posibil să vorbim despre fidelitatea definitivă a unui test: diferitele forme ale fidelității sunt deosebite în funcție de sursa erorii analizată într-un studiu particular. În plus, rezultatele studiului fidelității vor diferi în funcție de caracteristicile grupului studiat (omogenitatea referitoare la constructele măsurate are o puternică influență în calcularea coeficienților). În practică, coeficientul de fidelitate clasic indică, așa cum a fost menționat la itemul 4.2, măsura în care scorurile pot fi generalizate cu privire la formă, itemi, timp și evaluatori. Fidelitatea unui test este în general studiată folosind

metode clasice și itemii din acest capitol au fost explicați în consecință. În orice caz, este posibil să se folosească alte modele, de exemplu utilizarea analizei varianței pentru a estima influența diferitelor aspecte la un moment dat.

În general, este menționată evaluarea doar a unui singur coeficient de fidelitate, deși testul poate furniza diferite scoruri. De exemplu, acesta este cazul chestionarelor de personalitate și de interese care conțin diferite scale, și al bateriilor de aptitudini multiple care conțin câteva subteste (chiar dacă sunt administrate independent sau nu). Când scorurile la subteste sunt adunate pentru a obține un scor total, cum e cazul testelor de inteligență, există trei posibile presupuneri:

- Are importanță numai interpretarea scorului total. În această situație numai fidelitatea scorului total este nevoie să fie evaluată.
- Autorul testului afirmă că scorul total este într-adevăr cel mai important, dar că interpretarea scorurilor la subteste este și ea posibilă. În această situație fidelitatea scorurilor la subteste ar trebui evaluată cu un criteriu care are un nivel aflat imediat sub nivelul criteriului aplicat pentru scorul total. De exemplu, dacă scorul total este categorizat ca fiind “important”, scorul pentru subteste ar trebui să fie într-o categorie “mai puțin important”. În majoritatea cazurilor, scorurile la subteste sunt mai puțin fidele decât scorul total, dar când este aplicată regula de mai sus, evaluările pot fi aceleași.
- Autorul testului menționează că nu este nici o diferență a importanței, între scorurile la subteste și scorul total. În acest caz, cerințele pentru ambele tipuri de scoruri sunt aceleași.

Când evaluările date pentru coeficienții de fidelitate ale scorurilor la subteste și scorul total diferă, acestea ar trebui menționate într-o notă de subsol, adăugată la evaluare. Este de asemenea important să se noteze că numai o singură evaluare este dată când autorul unui test furnizează coeficienți de fidelitate pentru mai multe grupuri.

Explicații / recomandări pentru fiecare item

Itemul 4.1 Unele tipuri de coeficienți de fidelitate vor fi în general furnizați, dar și rezultatele studiilor de generalizare pot fi luate în considerare.

Itemul 4.2 Nu pot fi făcute afirmații generale despre nivelul dorit al coeficientului de fidelitate, cu toate că un test care este folosit pentru decizii importante este recomandat să aibă o fidelitate cât mai ridicată.

Itemul 4.2.a Consistența internă a diferitelor forme ale testului (forme paralele, de exemplu) poate fi determinată prin tehnici de corelație prin compararea între forme și conținutului itemului și (în mod clasic) prin compararea dificultăților, mediilor și varianței itemului. Corelația este o evaluare a fidelității ambelor forme ale testului. Fidelitatea testelor paralele poate fi folosită în cazul testelor cu timp impus. Corelația dintre două jumătăți ale testului, formate fie pe baza înjumătățirii timpului de testare sau împărțirea conținutului testului, poate fi considerată ca și fidelitate a testelor paralele. Ulterior poate fi aplicată o corectare a lungimii testului.

Itemul 4.2.b. Consistența internă referitoare la itemii (sau grupurile de itemi) din cadrul unui test este de obicei calculată cu ajutorul coeficientului Alfa – Cronbach. Deoarece coeficientul α este raportat la numărul itemilor, este important să conștientizăm că un număr foarte mare de itemi poate duce la un coeficient de fidelitate crescut, și în cazurile în care

corelațiile inter-itemi sunt moderate. Din perspectiva zilelor noastre, este nerecomandabil să se folosească clasicii coeficienți split-half (de înjumătățire), deoarece rezultatele depind de împărțirea arbitrară a itemilor în jumătățile testului. Coeficientul de consistență internă sau indicele de omogenitate nu sunt utili în cazul testelor de viteză. În asemenea situații, alte metode de calculare a fidelității se dovedesc a fi utile precum metoda testelor paralele sau metoda test-retest pot fi folosite. Cu toate acestea, multe teste de abilitate au, de asemenea, o limită de timp. Mai ales în cazul în care un procentaj semnificativ din subiecții testați nu au fost capabili să completeze ultimii itemi ai testului, consistența internă nu ar trebui calculată automat, deoarece poate rezulta o fidelitate care poate fi supraestimată. În această situație, o estimare rezonabilă a fidelității poate fi obținută prin: împărțirea testului în două jumătăți (de exemplu, itemii pari și itemii impari), administrarea subtestelor în timpul de testare înjumătățit, calcularea corelației dintre cele două jumătăți și, în final, corectarea corelației luându-se în considerare înjumătățirea lungimii testului. Când viteza nu este un factor principal (aceasta este atunci când cel puțin 70% din subiecții testați completează ultimul item), poate fi aplicată o formulă de corectare a coeficientului de consistență internă. Pentru toate tipurile de teste adaptive sau ajustate, consistența internă nu poate fi stabilită fără a se face o diferențiere. În acest caz, fidelitatea este estimată, folosindu-se regulile diferitelor puncte de întrerupere și prin calcularea corelațiilor dintre scorurile obținute și variabila criteriu.

Itemul 4.2.c. Consistența internă în timp este estimată prin metode de corelație test-retest care sunt calculate pentru administrări repetate ale testului la același grup. Trebuie fixat intervalul de timp dintre două aplicări, precum și lista evenimente relevante care au avut loc în acest timp. Dacă intervalul de timp este suficient de lung, el poate servi ca un semn al stabilității scorurilor la test.

Itemul 4.2.d Pentru metoda observației și scalele de evaluare în particular, este important să se știe dacă scorurile pot fi generalizate din punctul de vedere al observatorilor sau evaluatorilor. Indicii care pot fi folosiți sunt coeficienții de consistență precum: Cohen's kappa, coeficientul lui Gower (Gower, 1971) sau alte măsurări care iau în considerare diferențele dintre medii și varianțele evaluărilor.

Itemul 4.3.a Pentru fiecare din cele patru forme ale fidelității, unele puncte cu o importanță deosebită sunt redată mai jos:

1. Perfecționarea testelor sau scalelor, de obicei, își propune să obțină consistența internă cea mai ridicată. Aceasta poate rezulta într-un conținut al testului foarte specific care măsoară un construct foarte limitat. În general, nu este foarte dificil să se obțină o consistență internă crescută prin dezvoltarea itemilor care sunt aproape identici, dar asemenea scale sau teste s-ar putea să nu fie foarte utile. Aceste intercorelații relativ crescute pot fi datorate faptului că au o varianță neintenționată care nu este comună cu alți itemi din test. Aceasta se întâmplă atunci când itemii sunt formulați într-o manieră similară, au un anumit cuvânt în comun ș.a.m.d. Astfel de variante neintenționate pot contribui la o consistență internă crescută. În general vorbind, punctele menționate mai sus semnifică faptul că testul măsoară doar parțial un construct, altul decât cel stabilit (prin limitarea constructului sau prin introducerea de varianță neintenționată) pentru a se obține un grad crescut de omogenitate. Acest aspect poate fi evitat prin testarea unei singure dimensiuni în timpul fazei de perfecționare a testului, folosindu-se analiza LISREL, de exemplu.

2. Nu pot fi date standarde stricte în privința lungimii intervalului test-retest. Ca o regulă, un interval de timp foarte scurt (până la câteva săptămâni) nu este adecvat datorită rolului jucat de memorie. Un interval lung de timp (mai lung de un an) poate de asemenea, să nu fie folositor, deoarece evenimente externe sau experiențe pot influența personalitatea și abilitățile individului, afectând implicit și scorul la retestare. Corelația test-retest cu intervale lungi de timp nu este un indicator al fidelității testui, ci mai degrabă o măsurare a stabilității trăsăturii testate. În orice caz, intervalele menționate mai sus sunt mai degrabă stabilite arbitrar. Vârsta grupului testat, natura testului în sine și scopul testului trebuie întotdeauna luate în considerare, înainte de determinarea intervalului potrivit.
3. Metoda observației și a evaluărilor trebuie să fie realizate independent când fidelitatea inter-evaluării este folosită pentru a estima fidelitatea testului. Acest fapt ar trebui să fie clar descris în desig-ul cercetării.

Itemul 4.3.b. Coeficienții de fidelitate trebuie să fie calculați pentru grupurile pentru care testul este folosit. Acesta implică faptul că trebuie să fie calculați pentru fiecare grup etalonat, din moment ce scorurile subiecților testați sunt comparate cu un astfel de grup și contează fidelitatea obținută în cadrul acestui grup de referință. Din moment ce mărimea coeficientului de fidelitate depinde de distribuția scorurilor, coeficientul calculat va fi cu siguranță mai mare când scorurile întregului grup sau extremele distribuției scorurilor sunt folosite în locul coeficientului separat pentru fiecare grup etalonat.

Itemul 4.3.c. Mai jos sunt câteva exemple de informații care trebuie să fie valabile pentru a facilita sensul estimării calității pe care o are studiul fidelității:

- Deviațiile standard ale scorurilor de la testarea și retestarea grupului sunt date?
- Pentru testele cu limită de timp, a fost menționat procentajul subiecților care au răspuns la fiecare item?
- Eșantioanele pentru care au fost calculați coeficienții de fidelitate au fost descrise în detaliu?
- A fost menționat numărul observatorilor și evaluatorilor implicați în măsurarea coeficientului de fidelitate?
- Observatorii și evaluatorii, de obicei, sunt instruiți pentru sarcina lor. Acest training va influența calitatea evaluărilor și implicit nivelul fidelității inter-evaluării. Descrierea programului de instruire ar trebui detaliată suficient de bine ca să dea posibilitatea utilizatorilor de test să se pregătească în același mod încât fidelitatea evaluărilor să fie generalizată și în cazul lor. Ar trebui să fie realizabil pentru noii utilizatori să dobândească același nivel al acestei abilități. Este de asemenea important să se menționeze dacă coeficientul de fidelitate publicat este asociat cu estimarea unui singur observator sau evaluator, sau cu media estimărilor mai multor observatori sau evaluatori.

În cazul extrem, în care nu sunt date informații descriptive de nici un fel, coeficientul de fidelitate raportat poate fi evaluat cu “insuficient” deoarece calitatea design-ului cercetării nu poate fi verificat. În majoritatea cazurilor vor fi furnizate suficiente informații pentru a se cunoaște calitatea cercetării fidelității. În cazurile limită, în special (insuficient/suficient, suficient/bine), informații inadecvate pot constitui motiv pentru a se acorda evaluări scăzute.

5. Validitatea

Validitatea este măsura în care testul își atinge scopul. Pot fi conturate concluziile adecvate pornind de la scorurile testului? Acest sistem de evaluare urmărește clasificarea clasică în trei categorii cu privire la scopul cercetării validității (APA, 1985; Evers et al., 1988). Aceste categorii sunt: validitatea de construct, validitatea relativă la criteriu și validitatea de conținut. Validitatea de conținut a fost deja tratată la criteriul 1, deoarece este privită ca o parte din procesul de perfecționare a testului. Validitatea de construct se referă la răspunsul de la întrebarea: “Ce este măsurat cu ajutorul testului?”, “Testul măsoară conceptul propus sau parțial / în întregime măsoară altceva?”. Metodele și tehnicile utilizate frecvent pentru a dovedi caracterul evident al validității de construct sunt: analiza factorială pentru a demonstra unidimensionalitatea, compararea mediei scorurilor la grupurile care sunt planificate să difere, și calcularea corelațiilor dintre testele care se presupune că măsoară același construct. În mod esențial, acest tip de cercetare este ușor de realizat și rezultatele pot furniza un indiciu inițial al validității de construct (teoretice). Cu toate acestea, nici unul dintre acești indicatori nu este suficient pentru a garanta o evaluare de “suficient”. Numai acumularea unor dovezi privind un studiu mai extins al constructului sau un studiu bine proiectat cu privire la multi-caracteristici ale constructului pot duce la o evaluare de “suficient” sau “bine”.

Validitatea relativă la criteriu demonstrează că scorurile la test sunt raportate sistematic la unul sau mai multe criterii. În acest context, este utilizat termenul de predicție. Predicția poate fi focalizată asupra viitorului (validitate predictivă), asupra aceluiași moment din timp (validitatea concurrentă), sau asupra trecutului (validitate retrospectivă). Este important de menționat tipul criteriului pentru care sunt anticipate relațiile. Aceasta este adevărat, în special, când testul constă în mai multe subteste sau scale. Totuși, demonstrând validitatea tuturor subtestelor sau scalelor nu este esențial pentru a obține o evaluare “suficient” sau “bine”, dacă o singură scală validă poate face instrumentul util.

Explicații / recomandări pentru fiecare item (Criteriul 5A)

Itemul 5.1 Informațiile cerute aici se referă la structura internă și externă a testului. Structura internă poate fi investigată prin determinarea măsurilor asociate între (grupuri de) itemi, între itemi și test și între subteste. Proceduri precum a folosi subiecți la testare care gândesc cu voce tare când examinează itemii pot fi de asemenea utilizate. Structura externă este de obicei investigată prin determinarea relației cu alte teste (validitate convergentă sau divergentă).

Itemul 5.2 După cum a fost menționat mai sus, validitatea de construct este un aspect adăugat ca dovadă a cercetării. Studiarea validității de construct nu este niciodată completă. Cercetarea bias-ului referitor la itemi pentru grupuri diferite poate fi de dorit. Un avantaj în plus al cercetării bias-urilor în privința itemilor este că furnizează informații privind posibilitatea multidimensionalității constructului măsurat.

Itemul 5.3.a. Când rezultatele așteptate nu au fost formulate anticipat, rezultatele cercetării ar trebui interpretate cu o mare precauție. Fără așteptări clarificate, interpretarea rezultatelor studiului poate fi ușor degenerată într-o “pescuire”. Unele dintre corelațiile semnificative vor fi de fapt rodul întâmplării și nu este posibil să se vadă care sunt semnificative și care nu.

Itemul 5.3.b. Vezi explicațiile de la itemul 5.6.b.

Itemul 5.3.c. Fidelitatea măsurătorilor folosite trebuie să fie cunoscută. Va fi evident că validarea scorului unui test cu o fidelitate scăzută nu este folositoare, deoarece, rezultatele vor fi îndoielnice (o corelație scăzută poate semnifica faptul că testul măsoară ceva diferit sau poate fi rezultatul unei fidelități scăzute). Mai mult chiar, a valida un test cu unul asemănător este util numai dacă validitatea celui alt test a fost suficient investigată.

Itemul 5.3.d. Trebuie furnizate informații asupra mărimii eșantionului și o descriere suficient de detaliată a tehnicilor de analiză.

Explicații / recomandări pentru fiecare item (Criteriul 5B)

Itemul 5.4 Exemple vizând criteriile de raportare în privința validității predictive includ: corelația dintre scorurile la un test de inteligență și performanța școlară, indicele de predicție la un test folosit pentru selecția de personal și datele asupra sensibilității și specificității testelor pentru un diagnostic clinic. Aceste tipuri de date nu este necesar să fie adunate pentru fiecare test în fiecare situație nouă. Poate fi folosit principiul generalizării validității.

Itemul 5.5 Dacă este suficient unul sau mai mulți coeficienți de validitate, depinde de un număr de factori. Elementele cheie includ scopul testului, valoarea coeficienților de validitate, intervalul de încredere al coeficienților, valoarea testului comparabil cu alte instrumente sau alte surse de informație, corelația selectării și analiza cost-beneficiu. Cu cât este mai explicit autorul testului în privința scopului testului, cu atât mai bine evaluatorul va judeca dacă testul are o contribuție utilă la scopul propus.

Itemul 5.6.a Unele aspecte care trebuie luate în considerare sunt:

- Scorul predictor și scorul criteriu sunt stabiliți independent? De exemplu, acesta nu este cazul când supervisor-ul care evaluează performanța unui angajat știe rezultatele de la test dinainte de procesul de selectare.
- Intervalul de timp dintre administrarea testului și examinarea criteriului compatibil cu folosirea propusă a testului?
- Condițiile de testare din studiul de validitate sunt asemănătoare cu condițiile în care testul este în realitate folosit?
- Când este folosită generalizarea validității, autorul testului expune suficiente argumente pentru similaritatea dintre situații (sau teste) care sunt obiectul generalizării? Pentru a demonstra similaritatea testelor în sensul generalizării rezultatelor validității, trebuie arătat că ambele teste măsoară constructe similare și au fidelități identice.

Itemul 5.6.b. Cercetarea validității ar trebui realizată pe o populație pentru care testul este proiectat. Este cunoscut că indicii validității pot să descrească foarte mult când este folosit un grup omogen în locul unui grup eterogen (ținându-se seama de variabila măsurată). De exemplu, ar fi incorect să se valideze un test proiectat pentru selectarea terapiei pe un eșantion general de populație. Eșantionul trebuie să fie descris prin semnificația variabilelor relevante psihologic și demografic pentru a facilita o evaluare adecvată a acestui item.

Itemul 5.6.c. Câteodată alegerea unui criteriu este evidentă și scorurile criteriu pot fi ușor adunate, în timp ce în alte situații măsurile criteriu trebuie să fie în mod special

perfecționate și colectate. În toate situațiile criteriul trebuie decriptat în întregime și ar trebui indicate care aspecte comportamentale relevante sunt incluse în măsura criteriului și care nu. Aceasta este în special adevărat în cazul criteriilor mixte. Când intercorelațiile componentelor separate ale unui criteriu sunt scăzute, este mai bine să se determine coeficienții de validitate în mod separat pentru fiecare componentă decât doar pentru criteriul total.

Itemul 5.6.d. Când se evaluează calitatea studiului validității, ar trebui să fie disponibile câteva informații, care sunt expuse mai jos:

- Sunt oferite rezultatele obținute prin contra-validare?
- Dimensiunea eșantionului ar trebui menționată. Cu cât acesta este mai mic, cu atât mai largi sunt intervalele de încredere ale coeficienților de validitate.

Acest sistem de evaluare este o adaptare, la specificul românesc, a modelului olandez construit de Arne Evers.