

5. CATEVA DIN PROPRIETATILE PSIHOMETRICE ALE INSTRUMENTELOR STANDARDIZATE DE EVALUARE

Cu totii avem nevoie de a lua decizii cu privire la ingrijirea sanatatii. Complexitatea informatiei si dificultatea alegerilor sunt evidente si greu de depasit, in functie de pozitia ocupata in sistemul de ingrijire a sanatatii, ca clinet (pacient, consummator), furnizor de servicii de sanatate sau creator de politici de sanatate. Un client trebuie sa decida caror servicii, institutii sau agentii se adreseaza, in functie de intelegerea lui, asteptarile preformate, comfortul subiectiv oferit, accesibilitatea lui, etc. Furnizorul de servicii de sanatate trebuie sa decida ce fel de servicii ofera. Au ele vre-un rezultat (sunt effective)? Pot fi ele vandute (decontate) companiilor de asigurari de sanatate? Pot ele concura cu succes alte oferte de pe piata serviciilor de sanatate? Practicile oferite sunt suficient de confidente pentru a putea fi argumentate in instante juridice sau de alt fel atunci cand clientul le pune in discutie? etc. Nici creatorul de politici de sanatate nu este scutit de dileme de cunatificare cand trebuie sa decida caror nevoi comunitare trebuie sa raspunda programele de sanatate pe care le deseneaza si le implementeaza, sunt ele eficiente si care sunt indicatorii lor de evaluare?

Desi natura acestor decizii este diferita, ele toate se bazeaza pe tehnicile de masurare si decelare a fenomenelor procesului de sanatate si boala la un individ, grup sau comunitate. In fapt, atat in rutina clinica cat si in cercetare se raspunde la cele cinci intrebari cheie propuse de Dunn si Everitt, (1995)¹

Despre ce este vorba?	Problema de diagnostic
Cat de rea este?	Problema de masurare
Cat de comuna este?	Problema de prevalenta
Ce a cauzat-o?	Cautarea asociatiilor
Cum poate fi tratata?	Evaluarea terapiilor si rezultatului

Trecand in revista literatura, Fromberg si Kane (1989)² gasesc o larga controversa privitor la strategia de masurare a starii de sanatate. Ei considera ca exista cinci probleme nerezolvate:

- care sunt dimensiunile relevante ale sanatatii?
- cum trebuie sa fie prezentata de respondent starea de sanatate (de ex. trebuie sa evalueze fiecare dimensiune separat sau sa o evalueze global)?
- Care metoda de evaluare trebuie sa fie preferata (interviul standard, scala de evaluare, etc.)?
- Grupurile populationale difera in preferintele lor (de ex. populatia generala, profesionistii, pacientii)?
- Cum pot fi controlate variabilele situationale in asa fel incat masurarea sa fie acurata si confidenta?

De fapt este vorba de a alege dimensiunile de masurat si strategiile de masurare.

Dar se pune intrebarea: de ce atatea precautii? De ce trebuie sa masuram atat de exact? Un raspuns simplu dar si corect ar fi acela ca o masurare trebuie sa asigure *reproductibilitatea datelor*, respectiv siguranta sau increderea ca fenomenul supus masurarii ar da aceleasi rezultate la masurare daca aceasta ar fi facuta de diversi evaluatori dar cu acelasi instrument de masura sau la intervale scurte de timp. Conceptul de

reproductibilitate aduce certitudinea evaluarii, diagnosticului si deciziilor asteptate.

In termeni simpli, principalele metode de masurare in domeniul alcoolismului, metode bazate pe aprecieri subiective sau pe observarea comportamentului, se sprijina pe *scalarea* fenomenului supus masurarii, adica pe identificarea locarii acestuia pe un continuum care cuprinde teoretic toate starile sau situatiile in care se poate afla fenomenul sau subiectul. Prin scalare se obtine *scala de evaluare sau masurare ("rating scale")* care consta dintr-o linie imaginara cu capete si ancore bine definite. A aplica o scala de evaluare unui subiect inseamna a solicita respondentului sa se plaseze pe acest continuum utilizand o "ancora", respectiv punctul in care se opreste subiectul in evaluare sa. Scalele de evaluare sunt instrumentele de masurare cele mai comune utilizate in screening, diagnostic sau pentru evaluarea severitatii sau schimbarii unui simptom sau a unei trasaturi.

Mai exista si un alt tip de masurare, bazat pe culegerea standardizata a opiniilor, plangerilor, simptomelor, sentimentelor subiectilor prin chestionare sau interviuri. Aici este vorba mai curand de a *evalua* global un fenomen sau subiect fara a exista o unitate sau magnitudine a masurarii; este vorba de interviurile si chestionarele structurate de evaluare. Evaluarea poate fi definita ca un proces de aplicare a unei metode sistematice si structurate de descriere a unui fenomen sau trasaturi (Salvador-Carulla, 1996)³. Gradul de evaluare poate varia in functie de gradul de structurare a modului de culegere a fenomenului supus evaluarii, numit si algoritmul de cunatificare. Astfel, exista instrumente de evaluare semi-structurate sau structurate pe de-antregul. Ele se folosesc mai ales in evaluarea diagnostica sau a personalitatii.

Se considera ca prima scala de evaluare a fost conceputa de Walter D. Scott in 1920, presedintele unei societati de inginerie industrială din Chicago, care a vrut sa aibe un mijloc de evaluare, simplu si obiectiv, a calitatilor si defectelor candidatilor la un post de director. Se pare ca atunci s-a nascut si expresia atat de consacrata de "*rating scale*". Din alta perspectiva, prima incercare de a construi o scala de masurare entru scopuri clinice a fost facuta de Heymans si Wiersma (1909)⁴ care au redactat un inventar de simptome pe baza unui instrument de auto-raportare a acestor simptome. In 1950 apare prima scala folosita in psihiatrie, ce se datoreaza psihologului Wittenborn care a construit o scala de evaluare a psihopatologiei legata de evaluarea medicamentelor psihotrope (Wittenborn, 1950)⁵. Evolutia instrumentelor de evaluare a fost foarte rapida dupa introducerea medicamentelor psihotrope, aceste instrumente fiind necesare sa masoare efectivitatea terapeutica a acestor agenti si sa determine lansarea lor pe o piata cin ce in ce mai concurentiala. Cu toate acestea, noua viziune s-a impus greu si relativ tarziu, astfel ca in prima parte anilor '80, instrumentele standardizate nu erau mentionate in cele mai importante manuale sau tratate din SUA si Europa (Bobon, 1987)⁶. Am putea spune ca "dreptul la cetate" a acestui nou mod de abordare in psihiatrie si psihologie clinica nu este mai vechi de 25-30 ani. Pe aceste baze, Organizatia Mondiala a Sanatatii a intreprins eforturi constante si a antrenat intreaga comunitate stiintifica intr-un efort conjugat de standardizare, clasificare si statistica psihiatrica cu scopul dezvoltarii "*unui limbaj comun*"(Sartorius, 1983; Sartorius, 1989)^{7,8} care sa permita profesionistilor din domeniul sanatatii mentale si persoanelor implicate in managementul sanatatii mentale si problemelor psihosociale sa se inteleaga unii pe altii si sa lucreze impreuna.

Aceasta abordare este o noua provocare adusa stiintelor antropologice, care in deceniile viitoare va trebui sa aduca perfectionari continue ale procesului de evaluare si masurare a unor caracteristici umane, ele insele aflate in evolutie. Astfel, cuvintele lui Popper (1972)⁹ apar foarte actuale: "Stiinta apare ca o trecere de la arbitrar la conventional, dar la un conventional *controlat si ajustat*".

In psihiatrie, psihopatologie, psihologie clinica, psihoterapie, asistenta sociala si alte domenii ale sanatatii mentale, informatia este predominant descriptiva si scopul principal al instrumentelor standardizate de evaluare este de a *converti* aceste descriptii in date cunoscabile care sa fie susceptibile analizei statistice (Dennis si colab., 1995)¹⁰. Acest lucru nu este totdeauna posibil si nici chiar de dorit. Materialul scris sau verbal cules in exercitarea acestor domenii de activitate nu poate totdeauna si nici nu trebuie inlocuit cu evaluarile numerice. Ele sunt valabile mai ales pentru *asa-zisa masurare a schimbarii*. Ori de cate ori in starea de sanatate sau boala a unui subiect sau a unei populatii apare o schimbare, ea trebuie cuantificata cat mai exact pentru a putea fi pusa in relatie cu factorii determinanti, cu metodele curative, conturarea si dimensionarea metodelor profilactice si a politicilor comunitare in domeniu. In esenta, masurarea schimbarii este necesara pentru evaluarea *efectivitatii* metodelor de combatere si prevenire puse in opera si pentru desenarea balantei cost/beneficiu. De aceea, in ultimii ani s-a pus accentul pe proiectarea si realizarea de metode de evaluare cu o cat mai mare sensibilitate si specificitate, capabile sa deinceze cele mai mici schimbari ale fenomenului studiat.

In functie de domeniul folosit si alcoolismul nu face exceptie, instrumentele de evaluare se folosesc in patru scopuri importante: i) screening (detectia unui fenomen intr-o populatie), ii) diagnostic si

clasificare, iii) aprecierea importante/severitatii unui fenomen (simptom, opinie, emotie, comportament, etc) si iv) evaluarea rezultatului unui demers (pe termen scurt, mediu sau lung si rata succesului sau a reaparitiei/recaderii).

In paragrafele urmatoare vor fi expuse bazele procesului de masurare si evaluare in sanatatea mentala care va usura intelegerea modului de administrare, cotare, interpretare a instrumentelor care vor fi descrise in aceasta carte.

5.1. SCALELE DE EVALUARE

O scala de evaluare consta intr-o suita de itemi individuali, fiecare din ei acoperind un fenomen bine definit (de ex. simptom sau trasatura de personalitate) , care este evaluat conform gradului de severitate (Bech si colab, 1993)¹¹. In general, subiectul se pozitioneaza de-a lungul unei dimensiuni fizice sau cvasi-fizice a fenomenului, astfel incat anumite proprietati matematice sa poata fi deduse. In felul acesta, poate exista o locatie zero sau una de origine si intervale egale de evaluare, pe care un subiect se plaseaza ("ancoreaza") prin auto-evaluare. Prin scalarea cumulativa a fiecarui item se obtine un scor general la sfarsitul evaluarii. Aceasta natura cumulativa a evaluarii pentru fiecare item deosebeste scalele de evaluare de chestionare sau de inventarele de simptome sau opinii.

Scopul unei scale va determina continutul in itemi si diferitele aspecte ale structurii si aplicarii ei.

5.1.1. Constructia scalei

5.1.1.1. Sursa si formularea itemilor: Primul pas in redactarea unei scale este formularea sau inventarea itemilor ei. In general, itemii sunt

formulati pornind din cinci surse diferite: i) pacientii sau subiectii insisi, ii) observatia clinica, iii) teoria in domeniu, iv) datele furnizate de cercetare si v) opinia expertilor (Streiner si Norman, 1995)¹².

Foarte frecvent, pacientii si subiectii pasibili a fi cercetati prin scala ce se va dezvolta sunt o excelenta sursa a itemilor. Au fost dezvoltate o varietate de tehnici prin care sa se extraga in mod sistematic si viguros opiniile lor si aceste metode apartin domeniului cercetarii calitative. Printre cele mai cunoscute si folosite metode de analiza calitativa sunt "Interviul informantilor cheie" si metoda "Focus group"¹³. Itemii din sectiunile de alcool si drog din instrumentele recomandate de OMS, precum CIDI, SCAN si AUDADIS au fost selectionati, redactati si verificati cross-cultural prin astfel de metode, studii la care a participat si autorul acestei carti^{14, 15}.

Observatia clinica este cea mai frecventa si de incredere sursa pentru formularea itemilor unei scale de evaluare. In fond, scalele sunt un mod simplu de a strange observatii clinice in mod sistematic si standardizat. Exista insa riscul ca clinicianul sa greseasca in capitalizarea experientei sale in formularea unei scale, fie prin alegerea inadecvata a fenomenelor supuse masurarii sau prin definirea lipsita de acuratete a cadrului clinic. De aceea, experienta clinica trebuie sa fie confruntata cu datele teoretice, ai cele furnizate de cercetarea stiintifica si cu opinia expertilor in domeniu. Astfel se poate explica de ce multe scale de evaluare mai noi au aceiasi itemi cu alte chestionare sau scale mai vechi, doar cu mici modificari, generatia a doua de scale formandu-se in mare parte prin revizuirea primei generatii de scale.

5.1.1.2. Selectia itemilor: Dupa formularea unui numar oarecare de itemi, urmeaza selectarea lor dupa criterii adecvate pentru a fi inclusi in noua scala de evaluare.

Schematic, conform lui Streiner și Norman (1995)¹⁹ și Salvador-Carulla (1996)⁹ selecția itemilor se face conform următoarelor principii:

- comprehensivitatea: limbajul folosit să fie acceptat și adecvat ambianței socio-culturale a subiecților cărora le este adresat;
- să nu conțină ambiguități de formulare, să nu fie vag sau să conțină termeni rar folosiți sau un jargon specific;
- itemul să fie simplu formulat, să nu conțină mai multe întrebări în același timp sau să solicite mai multe răspunsuri în același timp;
- să se evite pe cât posibil itemii cu formulări negative;
- formularea itemilor trebuie să fie cât mai scurtă posibil
- să aibă putere discriminativă, respectiv să poată distinge subiecții care se poziționează diferit pe "plaja" scalei;
- definiția itemilor trebuie să fie exhaustivă și mutual exclusivă (*criteriul lui Guilford*)¹⁸;
- formularea itemilor nu trebuie să fie ofensivă sau trivială.

5.1.1.3. Numarul itemilor scalei: Scalele pot cuprinde un item, ca în cazul scalelor globale (de ex. scalele de evaluare a bunăstării, Scala Evaluării Globale sau Global Assessment Scale-GAS, etc) sau din mai mulți itemi. Există versiuni scurte sau lungi ale aceleiași scale (de ex. scala de depresie Hamilton, versiunea cu 21 itemi și cea cu 17 itemi). Dacă o scala are mai mult de 30 întrebări, atunci se vorbește de un chestionar; unii autori le denumesc "inventare" sau "checklist", mai ales atunci când se evaluează doar prezența sau absența fenomenului. Există tendința ca scalele de generația a II-a să fie mai scurte și să aibă în jur de 10 întrebări. După cum spune Bech (1993)¹⁶: "*Problema de bază aici este că nici un item al unei*

scale nu este suficient singur și ca un număr limitat de itemi luați împreună sunt suficienți dacă reprezintă universul problemelor ce trebuie investigate".

5.1.1.4. Scalarea răspunsurilor: Având desemnate un număr de întrebări după metodele descrise mai sus, se pune apoi problema de a alege o metodă prin care răspunsurile să fie obținute. Alegerea metodei este dictată de ce fel de întrebări am formulat și ce vrem să măsurăm prin ele. Astfel, se pot măsura *variabile categoriale*, precum statutul marital, statutul socio-economic, religia, etc. sau *variabile continue*, precum pentru măsurarea durerii, depresiei, presiunii sânghine, etc.

Pentru variabilele categoriale sunt construite scale în care se cere respondentului să furnizeze un răspuns "da - nu" (ca în cazul scalei MAST pentru screeningul alcoolismului), "adevărat - fals" (de ex. scala MacAndrew pentru screeningul alcoolismului) sau doar să bifeze un răspuns. Aceste scale se numesc *scale nominale de evaluare*.

Pentru variabilele continue există în mare trei feluri de măsurători: *tehnica estimării directe*, în care subiectul este invitat să indice răspunsul prin marcarea pe o linie sau bifarea unei casuțe; *metodele comparative*, în care subiectul alege dintr-o serie de alternative care au fost în prealabil calibrate; *metodele econometrice*, în care subiectul se evaluează prin ancorarea la stări extreme. Mai frecvente în domeniul de care ne ocupăm sunt metodele de estimare directă, precum *scalele analoge* și *scalele adjectivale*.

Scala analogă liniară vizuală (visual analogue scale - VAS) este o linie de o lungime de 70 până la 100 mm pe care sunt plasate diferite ancore. Subiectul este rugat să pună un "x" sau să trage o linie verticală în locul corespunzător cu starea percepută de el (fig. 1). Acest tip de scale permit subiectului să comunice ușor ce simte și o analiză parametrică a scorurilor

obtinute (Zeally & Aitken, 1969)¹⁷. Cu toate acestea, nu este un acord unanim asupra felului variabilelor ce le furnizeaza aceste scale: intervale sau ratii. Simplitatea acestor scale le-a facut foarte populare desi se considera ca in jur de 7% din subiecti nu sunt capabili sa le completeze.

Va rugam sa indicati printr-un semn pe scala de mai jos intensitatea tristetii pe care o simtiti

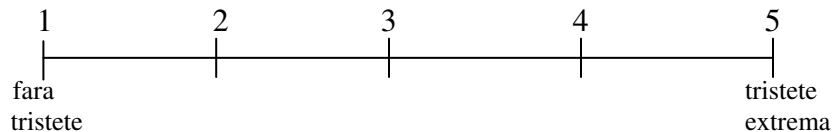


Fig. 1: Exemplu ipotetic de scala analoga vizuala de evaluare a tristetii

Scala analoga numerica este similara cu cea lineara cu exceptia ca dreapta este inlocuita cu numere (de la 0 la 7 sau 10 sau 100). Un exemplu este Scala globala de evaluare a functionarii (*Global Assessment of Functioning - GAF*)¹⁸ dar care este gradata de la 0 care inseamna fara functionare, la 100 care inseamna functionare foarte buna in toate ariile (social, professional, familial) si lipsa oricaror simptome.

Scalele adjectivale sunt mult mai raspandite in psihologie si psihopatologie si sunt focalizate pe descrierile adjectivale ale variabilelor continue sau categoriale. Evaluarea se face folosind categorii verbale sau adjective anterior calibrate care sunt folosite pentru a cuantifica severitatea unui fenomen. Ele pot cuprinde un singur item sau mai multi. Astfel de scale, cu alternative asezate gradual, se numesc si scale tip Likert. Se considera ca cea mai buna confidenta (*reliability*) o dau scalele cu 5 alternative, de ex. "intotdeauna", "deseori", "nici des, nici rar", "rareori", "niciodata". Un exemplu foarte cunoscut de scala adjectivala este scala numita SCL-90 (*Symptom Check List-90*) folosita pentru masurarea psihopatologiei generale si care este o scala Likert cu 5 ancore (de la 0 =

deloc la 4 = foarte mult)¹⁹. In aceasta scala, fiecare item este evaluat cu ajutorul scalarilor in 5 puncte (fig. 2).

In ultimul timp cat de mult ati fost deranjat de:

	Deloc	Putin	Moderat	Mult	Foarte mult
54. Lipsa de speranta	1	2	3	4	5

Fig. 2: Itemul No. 54 din SCL-90 si modul de scorare

Un alt tip de scala este asa numita *scala diferentiala semantica* (Osgood si colab. 1957)²⁰ care este folosita pentru a evalua un fenomen pe mai multe dimensiuni bipolare (fig. 3).

BOALA MEA ESTE:	
Dureroasa	Nedureroasa
Jenanta	Nejenanta
Serioasa	Usoara

Fig. 3. Scala diferentiala semantica (dupa Streiner & Norman, 1995)¹⁹

5.1.2. De la itemi la scala

Unele scale sunt constituite dintr-un singur item precum scalele analoge vizuale, altele din mai multi itemi desemnati sa evalueze o singura caracteristica de baza, precum Scala de Expectanta la Alcool a lui Young si Knight²¹ care contine 43 itemi, toti fiind cotate de la 1 la 6. Alegerea acestor itemi depinde de capacitatea lor de a "citi" fenomenul destinat a fi evaluat, respectiv pe baza relevantei clinice si validitatii lor. In plus, itemii alesi sufera o analiza de corelatie interna, in sensul ca fiecare dintre acestia trebuie sa coreleze pozitiv cu toti ceilalti si cu scala ca intreg in masura in

care toti sunt destinati sa evalueze acelasi fenomen. Prin analiza factoriala, itemii unei scale se pot grupa in factori in functie de capacitatea lor de a "incarca" acel factor.

In general, scorul unei scale se face adunand valoarea data de subiect fiecarui item in timpul procesului de evaluare, pornind de la premiza ca fiecare item are aceeasi importanta in cadrul scalei.

Exista si cazul in care in cadrul unei scale, itemii pot avea o "greutate" diferita in functie de ponderea pe care o au in evaluarea fenomenului. Un exemplu simplu este cel al Scalei de Dependenta la Alcool a lui Skinner si Allen²² in care fiecare item are o valoare diferita (de la 0 la 4) care se adauga la scorul final al scalei, valoare care este data de autorii scalei in functie de greutatea atribuita fiecarui item.

Atunci cand avem o distributie de scoruri intr-o populatie masurata cu o scala de evaluare a unui fenomen, uneori vrem sa predictam un rezultat dichotomic, cel mai ades de prezent sau absent. Astfel, daca masuram dependenta la alcool cu Scala de Dependenta a lui Skinner si vrem sa stim de la ce scor putem spune ca o persoana prezinta dependenta la alcool si de la care nu. Acest prag se numeste "prag de separare" (*cut-off point*) si se stabileste dupa un procedeu denumit *Receiver Operating Characteristic curve - ROC*^{23, 24}. Acest punct se calculeaza prin raportarea la doi indicatori statistici: *sensitivitatea* (rata cazurilor adevarat pozitive decelate de scala) si *specificitatea* (rata cazurilor adevarat negative decelate de scala). Un exemplu foarte cunoscut este instrumentul de screening pentru alcoolism - AUDIT (*Alcoholism Use Disorders Identification Test*) promovat de Organizatia Mondiala a Sanatatii (OMS)²⁵ la care punctele sau pragurile de separare au fost determinate prin metoda ROC: un scor mai mare de 8 identificand un consum de alcool daunator^{26, 27}.

Alte atribute ale scalelor de evaluare sunt normele legate de vârstă și sex. Pentru multe scale se furnizează de către autor norme diferențiate pe grupe de vârstă sau în funcție de sex, considerându-se că fenomenul măsurabil este diferit în funcție de aceste variabile.

5.1.3. Administrarea scalelor

Administrarea scalelor de evaluare se face, în general, în funcție de tipul de informații dorit a fi obținut și de sursa informațiilor. Astfel, scalele bazate pe interviul cu pacientul sau pe examinarea sa sunt denumite *scale de observare*. Scalele care se completează de către subiect, fără ajutorul unei alte persoane se numesc *scale de auto-evaluare*. Aici avem de fapt problema auto- sau hetero-evaluării. S-ar putea crede că scalele completate de un observator sau o terță persoană sunt superioare în ceea ce privește confidența datelor evaluate față de scalele completate de subiectul însuși și aceasta cu atât mai mult în cazul subiecților consumatori de alcool. Cu toate acestea, s-a observat în multiple trialuri clinice că auto-evaluarea își păstrează întreaga încredere, mai ales în evaluările repetate la intervale scurte sau la cele făcute pe micro-computer.

Pentru scalele de observare trebuie să se specifice pentru fiecare item sau pentru scala întreaga dacă se bazează pe situația de interviu (față în față) sau pe alte situații ca în observația simplă, fără conversație, în cazul comportamentului motor, sau pe observația mediata de echipamente diverse, de exemplu camere video, etc.

5.2. INTERVIURILE DE DIAGNOSTIC

La celalalt capat al continuumului instrumentelor de evaluare, la care la un alt capat se afla scalele de evaluare, se gasesc interviurile. In principal un interviu este o interactiune intre doi oameni in care o persoana intreaba (interviewatorul) si cealalta furnizeaza raspunsurile (interviewatul)²⁸. Dupa Zubin (1989)²⁹, exista patru criterii prin care un interviu diagnostic folosit in evaluarea sanatatii mentale se deosebeste de conversatie sau comunicare: i) intalnirea interlocutorilor se face fata in fata; ii) este directionat de scopul particular al interviewatorului; iii) foloseste concomitent conversatia si observatia si iv) implica o relatie ierarhica intre interlocutori, in sensul ca unul este interviewator si celalalt este interviewat.

In functie de modul de administrare a intrebarilor, exista *interviuri informale*, in care intrebarile sunt la latitudinea interviewatorului, atat ca formulare cat si ca secventa, si *interviuri standardizate*, in care formularea si secventa intrebarilor este stabilita ca regula stricta de creatorii interviului si la care cel ce il administreaza trebuie sa adere din convingere. Interviuurile informale au avantajul ca, prin flexibilitatea pe care o au, se muleaza mai bine pe relatia interviewat-interviewator, reusesc sa demonteze reticenta interviewatului si sunt expresia ideologiei si experientei interviewatorului, dar nu ofera certitudinea evaluarii, pentru ca repetat la acelasi subiect sau la subiecti diferiti rezultatele interviului pot varia nepermis de mult. Interviuurile standardizate au avantajul unei reproductibilitati maxime a rezultatelor, o confidenta marita asupra capacitatii de evaluare a fenomenului, dar ofera putina flexibilitate care sa permita adaptarea la conditiile subiectului sau ambiantei evaluarii.

In mare, scopurile interviului clinic psihiatric sunt: a) sa extraga informatii despre simptomele si antecedentele episodului curent; b) sa determine atitudinile si sentimentele pacientului in legatura cu simptomele

pe care le acuza și c) să se pună în lumină comportamentul nonverbal al interviuatului cu scopul de a se stabili natura problemelor evaluate³⁰.

Gradul de succes al unui interviu clinic este măsura în care el reușește să îndeplinească aceste deziderate. După Reiser (1984)³¹, interviurile diagnostice au două mari scopuri: a) să stabilească și să mențină o bună colaborare cu interviuatul și o bună complianță și b) să culegă informații care să permită formularea diagnosticului și planului terapeutic.

Indiferent cum ar fi, structurate sau nu, interviurile sunt o colecție de întrebări sau sugestii de întrebări care își au originea în concepția creatorilor lor asupra fenomenului de evaluat sau diagnosticat. Astfel, mai ales în cazul interviurilor standardizate, întrebările sunt formulate în funcție de "*definițiile operationale*" ale simptomelor sau elementelor care compun fenomenul de evaluat sau diagnosticat, definiții care asigură astfel un cadru strict, un referențial, care crește probabilitatea ca la fiecare aplicare, interviul să măsoare ceea ce și-a propus prin construcție și să împiedece distorsiunile în evaluare. Interviurile standardizate sunt bazate și pe o strictă formalizare a procesului de diagnostic prin care creatorul interviului își manifestă aderența la o anumită doctrină sau clasificare, în funcție de care construiește matricea interviului, respectiv domeniile și secțiunile lui, algoritmul întrebărilor, regulile de înregistrare a răspunsurilor și sistemul de codificare.

Superioritatea interviurilor structurate asupra interviurilor nestructurate a fost evidențiată de mai multe ori. Ele sunt folosite atât în studiile epidemiologice și clinice în vederea identificării "cazului" și a clasificării tulburărilor clinice, cât și în scop diagnostic³².

Primul interviu clinic structurat cu scop diagnostic (*Mental Status Schedule*) a fost conceput de Robert Spitzer, Joseph Fleiss și colaboratorii lor în 1964 în cadrul Institutului Psihiatric al Statului New York³³. Acest

instrument a generat apoi celebrul Inventar de diagnostic pentru tulburari afective si schizofrenie (*Schedule for Affective Disorders and Schizophrenia*)³⁴. Ambele interviuri diagnostice aveau o sectiune distincta pentru tulburarile mentale date de consumul de alcool. De aici inainte se deschide larg drumul aparitiei numeroaselor interviuri de diagnostic si a perfectionarii lor continue prin versiuni succesive.

Dupa cum am mentionat mai inainte, exista o serie de critici in legatura cu folosirea interviurilor structurate printre care am aminti aici faptul ca ele se abat prea tare de la principiile psihoterapeutice, sunt prea rigide, nu sunt economicoase ca timp si bani si induc un stil artificial al interviului³⁵. Cu toate acestea, Saghir (1971)³⁶ nu gaseste diferente intre interviurile structurate si cele nestructurate in ceea ce priveste capacitatea lor de a citi empatia subiectilor intervievati..

In ultimul deceniu s-a conturat ideea ca fiecare clasificare a tulburarilor mentale sa fie insotita si de instrumentul de diagnostic adecvat. Astfel Organizatia Mondiala a Sanatatii promoveaza Interviul diagnostic international compozit (*Composite International Diagnostic Interview - CIDI*)³⁷ si Inventarul pentru evaluarea clinica in neuropsihiatrie (*Schedules for Clinical Assessment in Neuropsychiatry - SCAN*)³⁸. Avand in vedere complexitatea mereu crescanda a acestor instrumente, aplicarea interviurilor clinice structurate necesita un antrenament specific si cunoasterea detaliata a modului de intervievare si cotare. Multi autori^{39, 40} au aratat ca tehnica interviurilor diagnostice trebuie sa reprezinte unul din obiectivele principale a educatiei in sanatate mentala. Bögels si colab. (1995)⁴¹ au demonstrat ca o cunoastere adecvata a procedurii specifice de administrare a unui interviu diagnostic duce la cresterea validitatii constructive a lui. Astfel, acesti autori deosebesc doua feluri de abilitati sau calificari legate de cunoasterea unui

interviu diagnostic: a) abilitatea procesuala (*process skill*) este dobândită prin cunoașterea procedurilor de administrare și cotare care contribuie la stabilirea unei bune relații intervievator-intervievat și b) abilitatea de conținut (*content skill*) care contribuie la o bună colectare a informațiilor diagnostice și la o adecvată formulare diagnostică. Această carte încearcă tocmai să îmbunătățească abilitatea persoanelor implicate în evaluarea și tratamentul alcoolismului.

5.3. PARAMETRII DE CALITATE A INSTRUMENTELOR DE EVALUARE

Există astăzi o multitudine de interviuri și mai ales de scale de evaluare. Practicianul și în special cercetătorul se află într-o pronunțată dilemă referitor la alegerea instrumentului necesar scopului propus. În literatura de specialitate există astăzi mai multe lucrări de referință care ghidează o astfel de alegere. Principalul criteriu după care se alege un instrument de evaluare ai se respinge un altul este calitatea lui. De obicei, judecarea acestei calități este făcută de însuși autorii lui și aceste norme de calitate sunt publicate odată cu instrumentul propus. Apoi, comunitatea științifică are bunul obicei de a testa în diferite circumstanțe instrumentul nou promovat și astfel, el trece sau nu de ultima încercare înainte de a fi adoptat și larg recomandat.

Cei mai cunoscuți parametri de calitate a instrumentelor de evaluare sunt validitatea și confidența (*reliability*).

5.3.1. Validitatea

Pentru a avea o normă de calitate a unui instrument de măsură trebuie să avem un etalon, un etalon absolut sau ceea ce se numește "*gold standard*"

(etalon de aur). In cazul procedurilor de diagnostic "*gold standard-ul*" este o proba de laborator, o analiza anatomo-patologica, o examinare radiologica, etc. care este indubitabila pentru catalogarea diagnostica: de exemplu, o determinare a glucozei in sange pentru diagnosticul de diabet. Este de la sine inteles ca conceptul de "*gold standard*" este o idealizare dar subliniaza tendinta medicinei de a cauta pentru fiecare boala probe de laborator de necontestat, in masura in care nici anamneza si nici examinarea fizica a pacientului nu este capabila sa puna un diagnostic de certitudine.

In diagnosticul tulburarilor mentale nu putem vorbi inca de "*gold standard*" si din acest motiv Spitzer (1983)⁴² introduce un alt etalon numit "*LEAD standard*" care este acronimul a trei concepte esentiale: evaluarea longitudinala (L), datele furnizate de experti (E) si folosirea tuturor datelor disponibile (AD - *all data*). "Longitudinal" se refera la procedura diagnostica care nu trebuie limitata la o singura examinare. Lungimea perioadei dintre examinari poate varia in functie de modalitatea de evolutie si prezentare a unei tulburari psihice. "Expert" indica ca diagnosticul trebuie facut de clinicieni experientati. "Toate datele" arata ca clinicienii experti trebuie sa folosesca nu numai evaluarea sistematica si longitudinala a subiectului, dar sa utilizeze si date de la alti informanti semnificativi precum membrii de familie, personal de ingrijire, terapisti, date clinice mai vechi sau teste de laborator. Standardul LEAD a fost folosit la evaluarea validitatii procedurale a mai multor instrumente de diagnostic precum Programul de Interviu Diagnostic (*Diagnostic Interview Schedule*)⁴³, a interviurilor de personalitate⁴⁴, sau a interviurilor pentru diagnosticul alcoolismului si abuzului de droguri⁴⁵.

Dar ce este validitatea? In teoria psihometrica, validitatea este gradul sau marimea cu care un instrument masoara ceea ce este presupus sa

masoare. Spre deosebire de masuratorile din sanatatea fizica, cele adresate sanatatii mentale sunt marimi subiective, traite diferit de fiecare individ. Cand vorbim de masurarea "calitatii vietii" sau a "suportului social", aceste evaluari depind de definitiile ce se acorda acestor variabile, altfel ca rezultatele masuratorilor vor diferi de la om la om sau de la un instrument la altul destinat sa le evalueze. Cand vorbim de un diagnostic psihiatric sau psihologic, trebuie sa mentionam criteriologia sau alte argumente pe care se sprijina. Validitatea instrumentelor de masura ofera garantia ca desi evaluam variabile subiective, rezultatele masuratorilor pot fi comparabile de la un subiect la altul. Deci tinta masurarii trebuie sa fie foarte bine stabilita si circumscrisa, altfel evaluarea validitatii unui instrument va fi foarte greu de apreciat. In fond, validitatea trebuie sa fie consistenta cu constructele teoretice ce au stat la baze constructiei unui instrument de masura.

Prin testarea validitatii unei scale sau interviu noi demonstram proprietatile psihometrice ale acelu instrument de a caracteriza subiectii care sunt evaluati. Cum Landy (1986)⁴⁶ spunea: "*procesul de validare nu este atat de directionat asupra integritatii testului cat este dirijat spre deductiile ce pot fi facute despre attributele subiectilor ce au produs scorul testului*". Cu alte cuvinte, validarea unei scale este un proces prin care determinam gradul de incredere asupra rezultatelor testului si despre deductiilor ce le facem despre subiectii ce au dat diferite scoruri la scale ce li s-a aplicat.

5.3.2. Senzitivitatea, specificitatea si acuratetea

Alta categorie de caracteristici de calitate ale instrumentelor de masurare este data de capacitatea unui scale sau interviu de a fi cat mai adecvat detectarii (nu numai masurarii) starii sau trasaturii pentru care a fost destinat. Aceasta capacitate se determina atunci cand avem un standard de

comparare, respectiv un "gold standard", un "LEAD standard" sau cel puțin un diagnostic facut de experti. Acesti indicatori de calitate se determina cel mai bine utilizand un tabel 2 x 2 precum este prezentat in tabelul 1.

ALCOOLISM		LEAD STANDARD	
		diagnostic	Non-diagnostic
Instrument de masura	Prezent	adevarat pozitiv (a)	fals pozitiv (b)
	Absent	fals negativ (c)	adevarat negativ (d)

Tabelul 1: Tabel 2 x 2 pentru calcularea senzitivitatii si specificitatii

Dupa cum se vede in tabel, subiectii "adevarat pozitivi" sunt cei care sunt decelati avand diagnosticul de alcoolism atat conform diagnosticului clinic cat si de instrumentul de evaluare, in timp ce subiectii "adevarati negativi" sunt cei care nu indeplinesc nici diagnosticul si nici rezultatul de la instrumentul de masurare. Subiectii care au fost decelati ca avand alcoolism de catre instrument dar nu au primit diagnosticul de alcoolism de la clinicieni sunt considerati "falsi pozitivi", iar cei care au fost gasiti indemni de alcoolism la test dar au primit diagnosticul de alcoolism de la clinicieni sunt "falsi negativi".

Formulele dupa care se calculeaza acesti parametri de calitate ai instrumentelor de masura sunt^{1,2}:

$$\text{Senzitivitatea} = \frac{a}{a + b} \times 100$$

senzitivitatea fiind capacitatea unui test de a identifica corect o boala sau un simptom la o persoana care o are intr-adevar.

$$\text{Specificitatea} = \frac{d}{c + d} \times 100$$

specificitatea fiind capacitatea unui test de a extrage numai bolile sau simptomele pe care trebuie sa le detecteze ai nu altele.

$$\text{Acuratetea} = \frac{a}{a + c} \times 100$$

acuratetea fiind increderea care poate fi atribuita unui rezultat al testului, o acuratete de peste 80% aratand ca in aceasta proportie au fost oferite raspunsuri pozitive la chestionar.

Asa cum am amintit mai sus, pe baza indicatorilor de senzitivitate si specificitate, unei scale i se pot stabili punctele de separare (*cut-off points*), respectiv de la ce scor un test distinge sau identifica un anumit fenomen. Metoda este prea laborioasa ca sa poata fi descrisa aici.

5.4. CONFIDENTA

Confidenta este un alt parametru major dupa care se judeca calitatea unui instrument de masura. El este traducerea cuvintului englezesc "*reliability*" (incredere, siguranta, soliditate, etc.) . Francezii il traduc prin "fidelitate". Unii cercetatori tind sa foloseasca o lista larga de sinonime printre care amintim: "obiectivitate", "reproductibilitate", "stabilitate", "acord", "asociatie", "senzitivitate", "precizie".

Se știe că orice măsurare duce inevitabil la unele erori din varii motive. Conceptul de confidență este un mod fundamental de a reflecta gradul de eroare inerent oricărei măsurări¹² sau gradul cu care rezultatul măsurătorii este reproductibil³. Noțiunea de "eroare a măsurării" recunoaște că scorul unui test și alte feluri de observații cantitative sunt subiect de fluctuație de la valoarea "ipotetic" adevărată care nu poate fi determinată în sine dar care este "estimată" prin teste sau observații⁴⁹. Cu cât această eroare este mai mică cu atât datele obținute sunt mai confidente. Măsurători independente care nu obțin rezultate asemănătoare sunt suspicinate de a avea erori într-una sau ambele măsurări. Pe de altă parte, chiar măsurările care dau rezultate foarte asemănătoare sau chiar exacte pot conține erori. De aici se poate vedea importanța evaluării confidenței instrumentelor folosite pentru măsurare.

Confidența măsurării aduce informații nu numai despre eroarea măsurării ci și despre variabilitatea subiecților supuși măsurării, de exemplu bolnavi, clienții serviciilor de sănătate, etc. Există o variabilitate așteptată și sistematică a rezultatelor subiecților la o măsurare dată, explicabilă prin "*adeverata*" diferență dintre subiecți, indiferent dacă ei suferă de aceeași boală sau li se măsoară aceeași trăsătură. Astfel, eroare de măsurare ar fi "*falsa*" diferență dintre subiecți. Coeficientul de confidență reflectă mărimea cu care un instrument de măsură poate diferenția subiecții¹².

Deși se crede că confidența este o măsură a unui test, trebuie să se spună că acest parametru este strâns legat de populația pe care se vrea să se aplice măsurarea. Coeficientul de confidență are înțeles doar atunci când este aplicat unei anumite populații și în anumite condiții de măsurare. Astfel, un anumit coeficient de confidență a unui instrument de măsură este o afirmație

despre cum măsura este reproductibilă în anumite condiții și la anumite populații.

Există mai multe feluri de coeficienți de confidență în funcție de tipul de instrument și de intenția experimentatorului. *Confidența internă* a unui instrument înseamnă soliditatea psihometrică a lui, structura internă și nivelul la care diferiții lui itemi sunt intercorelați³. O presupunere generală asupra unui test cu mai mulți itemi este că toți itemii sunt legați de același construct pe care se dorește să-l evalueze. În felul acesta se înțelege că toți itemii vor trebui să fie corelați unul cu celălalt ceea ce subliniază *omogenitatea* aceluși test.

În măsura în care confidență este un indice al reproductibilității unui test, *confidența externă* măsoară stabilitatea măsurării cu acel test în ciuda schimbărilor parametrilor externi sau condițiilor măsurării, adică a schimbărilor care nu sunt inerente aplicării testului. Acești factori externi pot fi: ambianța în care se aplică testul, maniera în care este aplicat de către intervievator, factori subiectivi care tin de relația ce se stabilește între intervievat și intervievator, etc. În esență, confidența externă se reduce la gradul de potrivire și acord ("*agreement*") între două măsurări.

Confidența externă include *confidența intra-observator* sau evaluator și *confidența inter-observator/evaluator*. Confidența intra-observator se referă la stabilitatea a două măsurări făcute de același evaluator în momente diferite și se mai numește și *confidența test-retest*. Confidența inter-observator este gradul de consens obținut între mai mulți observatori care au efectuat același test. Uneori se poate vorbi și de *confidența inter-informanți* atunci când un evaluator folosește mai mulți informanți, ca atunci când se pune diagnosticul de alcoolism folosindu-se informații din familie.

Confidenta test-retest se stabileste prin administrarea interviului sau scalei la un grup de subiecti in doua ocazii separate de o perioada specificata de timp, in general intre 3 la 14 zile. Confidenta se masoara prin coeficientul de "agreement" numit *kappa* (Cohen, 1960)⁵⁰ care se intinde de la -1 (dezacord total) la +1 (acord total). Un indice de 0,75 sau mai mult indica un acord excelent, intre 0,65 la 0,74 un acord bun, intre 0,40 si 0,64 acord multumitor si sub 0,40 acord prost (Fleiss, 1981)⁵¹. Coeficientul de confidenta *kappa* depinde in mare masura de prevalenta fenomenului studiat, respectiv cu cat prevalenta este mai mare cu atat increderea pe care putem sa o acordam coeficientului *kappa* calculat este mai mare.

Un coeficient de confidenta prost nu indica totdeauna faptul ca testul este prost. Astfel, evaluarea discrepantelor dintre prima evaluare si cea de a doua, procedeu promovata de Cottler si colab. (1994)⁵² numit Interviul discrepantelor, permite evidentierea surselor de neconfidenta care pot apartine, in mare, intervievatorului, respondentului sau procesului de evaluare in sine.

Pentru calcularea confidentei inter-observatori se poate utiliza la fel metoda test-retest dar fiecare evaluare va fi facuta separat de un alt evaluator, de obicei in perechi iar apoi se calculeaza coeficientul *kappa*.

O buna exemplificare a acestor metode o reprezinta recentul studiu OMS asupra validitatii si confidentei instrumentelor CIDI⁵¹, SCAN⁵² si AUDADIS (*Alcohol Use Disorder and Associated Disabilities Interview Schedule*)⁵³ in diagnosticul tulburarilor de consum de alcool si droguri, studiu la care am participat personal (Üstün si colab., 1997)¹⁴. In tabelul Nr. 2 se prezinta centrele participante, instrumentele, domeniile de baza ale variabilelor studiate ai nivelul de validitate si confidenta din acest studiu.

Centrele	Substantele	Instrumentele	Sistemul	Nivelul de acord	Nivelul validitatii si confidentei
Amsterdam	Alcool	AUDADIS	ICD-10	1. Item	1. Evaluarea independenta clinic (EIC)
Ankara	Cannabis	SCAN	- consum daunator	2. Criteriu	2. Diagnostic clinic
Atena	Amfetamine	CIDI	- dependenta	3. Diagnostic	3. Analiza in urina (ONTRAK)
Bangalore	Sedative		DSM-III-R	4. Interviu de discrepanta	4. Diagnostic clinic versus EIC
Farmington	Cocaina		- abuz		5. Test-retest, kappa
Ibadan	PCP		- dependenta		6. Comparatie intre teste
Jebel	Halucinogene		DSM-IV		
Luxemburg	Inhalanti		- abuz		
St.Louis	Altele		- dependenta		
San Juan			A. Orice diagnostic vs nici un diagnostic		
Sydney (1)			B. Grupe specifice de Diagnostic		
Sydney (2)					

Tabelul Nr. 2: Prezentare sumara a studiului OMS de validitate si confidenta a instrumentelor de diagnostic ale tulburarii uzului de alcool si droguri (Üstün si colab, 1997)¹⁴.

Acest studiu care a determinat validitatea si nivelul confidentei pentru aceste interviuri va fi prezentat mai pe larg in capitolele urmatoare, atunci cand se va discuta fiecare instrument in parte.

Tot in cadrul acestui studiu OMS s-a facut o comparare a confidentei AUDADIS-ului in diagnosticul tulburarii de consum a alcoolului si drogurilor intre trei esantioane de utilizatori din Australia, India si Romania, folosindu-se metoda test-retest si calcularea coeficientului de acord kappa (Chatterji, Saunders, Vrasti si colab. 1997)⁵⁴. Vrasti, Grant, Chatterji si colab.⁵⁵ publica in 1998 primul studiu de confidenta a unui versiuni ramanesti a unui instrument international promovat de OMS si Institutul National de Alcoolism al SUA, respectiv AUDADIS, folosind tehnica de evaluare a confidentei inter-interviewatori prin studiu test-retest si calcularea coeficientului kappa si a discrepantelor dupa metoda Cottler.

5.5. CONFIDENTA CROSS-CULTURALA

Factorii culturali joaca un rol major in toate aspectele evaluarii, diagnosticului si terapiei tulburarilor mentale si a celor date de uzul de

alcool si droguri. Ei influenteaza felul in care subiectul vede boala sa sau a altor, expresivitatea problemelor sale si influenteaza relatia pe care o are cu acei ce o evalueaza si o trateaza. Toate acestea se petrec pentru ca "*atat pacientul cat si doctorul sunt membrii ai traditiilor etnoculturale care decodeaza si socializeaza variate supozitii si premize privind sanatatea si boala si din cauza ca doctorul apartine unei subculturii medicale care , la fel, decodeaza si socializeaza variate supozitii si premize despre boala si sanatate*" (Marsella si Westermeyer, 1993)⁵⁶.

Unul din principalele motive pentru care factorii culturali joaca un rol important, dar adesea ignorat, in procesul de evaluare si diagnostic al problemelor de sanatate mentala ale indivizilor este *etnocentrismul* frecvent al sistemului de ingrijire a sanatatii mentale. Acest etnocentrism isi are radacinile in valorile, stereotipurile, credintele si practicile legate de viziunea si componentele cognitive ale definitiei sanatatii sau bolii mentale. El limiteaza semnificativ capacitatea profesionistilor si laicilor de a se alinia la datele stiintifice internationale si parca niciunde nu se vede mai pregnant aceasta ancorare in etnocentrism ca atunci cand se cauta sa se introduca si sa se utilizeze instrumente internationale de masurare si diagnostic in domeniul sanatatii mentale, lucru necunoscut in alte domenii ale medicinei, de exemplu.

Folosirea unui asemenea instrument international de evaluare sau diagnostic in diferite culturi implica o serie de probleme metodologice complexe a caror evitare duce la pierderea validitatii si confidentei lor initiale. Aceste probleme fac parte din ceea ce se cheama cercetarea "potrivirii cross-culturale" (*cross-cultural applicability research*) a instrumentelor utilizate. Parintii unui instrument, care se doreste a capata relevanta internationala, trebuie sa se ingrijeasca atat de validitatea si

confidenta lui dar si de aceasta "potrivire cross-culturala" care ar putea fi denumita si ca "confidenta cross-culturala", cu alte cuvinte daca se poate folosi cu aceiasi incredere in diferite culturi sau daca furnizeaza aceleasi scoruri sau date in diferite culturi la subiecti cu aceiasi tulburare mentala. Dupa Flaherty si colab. (1988)⁵⁷ propun o "echivalare" cross-culturala pe patru nivele a unui instrument international:

1. Echivalare continutului: daca fiecare item al instrumentului este relevant pentru fenomenul masurat in cultura data;
2. Echivalare lexicala si sintactica: verificarea daca intelesul fiecarui item in fiecare cultura este la fel;
3. Echivalare tehnica: daca metoda de administrare este adecvata si comparabila in fiecare cultura;
4. Echivalare conceptuala: daca constructele teoretice de la baza construirii instrumentului sunt valabile in fiecare cultura.

Toate aceste deziderate sunt pe deplin implinite prin metodologiile cunoscute ale anchetelor etnografice din studiile de antropologie culturala: traducere-retrotraducere (*translation – back-translation*), interviul informanților cheie si grupul focus (Trotter, 1991)⁵⁸. Astfel, studiile etnoculturale trebuie sa ia in considerare problema "*echivalentei*", respectiv a marimii cu care un cuvânt, propozitie, denumire, scala sau o norma a unui instrument de masura poate fi considerata similara sau echivalenta in mai multe culturi (Marsella si Kameoka, 1989)³. Facand aceasta echivalare inseamna a construi un pod intre conceptele profesionale din mai multe culturi, de fapt, a construi confidenta etnoculturala a procesului de masurare.

Un exemplu concludent si foarte recent este studiul OMS asupra potrivirii cross-culturale a diagnosticelor si evaluarii tulburarilor folosirii alcoolului si drogurilor (Room si colab. 1996)¹⁵ efectuat intre 1992-1996 in

9 centre din tot atatea tari, la care autorul a participat in calitate de investigator principal al centrului colaborativ de la Spitalul de Psihiatrie Jebel. Principalul scop al acestui studiu a fost sa se determina prin metodologiile mentionate mai sus, daca formularile diagnostice, criteriile diagnostice, simptomele si itemii folositi de instrumentele de evaluare (respectiv CIDI, SCAD si AUDADIS) au acelasi inteles in cele noua culturi diferite, atat pentru profesionisti cat si pentru publicul larg (laici) si de a gasi calea de a netezi si unifica toate discrepantele gasite. Foarte sugestiv, Philippe Batel, comentand rezultatele acestui studiu, spunea ca studiul a trebuit sa rezolve problema "construirii unui lift in turnul Babel" in care cele noua culturi sunt noua etaje iar liftul (conceptele OMS vehiculate de profesionisti) trebuie sa faca legatura si comunicarea dintre aceste etaje ale turnului Babel¹⁵.

5.6. ALEGEREA UNUI INSTRUMENT DE EVALUARE

Chiar daca avantajele folosirii instrumentelor standardizate sunt evidente, procesul de selectie a unui instrument de masura nu este totdeauna simplu. In selectarea acestuia exista mai multe criterii generale precum⁴:

1. Instrumentul trebuie sa produca informatii utile pentru evaluarea si diagnosticul problemelor subiectului/clientului, planificarea tratamentului, urmarirea progresului in timpul tratamentului si, daca este planificata si o perioada de urmarire, evaluarea subiectului o perioada anumita dupa tratament.
2. Instrumentul trebuie sa produca informatii care sunt cuantificabile si, daca este posibil, sa produca scoruri care pot fi folosite in compararea statutului clientului/subiectului de-a lungul timpului

sau cu alti subiecti. Astfel de scoruri sunt folositoare si pentru a facilita o analiza statistica a performantei unui grup de subiecti.

3. Instrumentul sa fie relativ usor de administrat si nu excesiv de lung.
4. Limbajul si cuvintele folosite in formularea itemilor sa fie apropiate de statutul intelectual si cultural al clientului/subiectului, sa fie acceptabile pentru el si sa nu fie ofensatoare.
5. Instrumentul sa fie suficient de senzitiv la problemele ridicate de client/subiect astfel ca schimbarile semnificative ale acestora sa poata fi evaluate.
6. Costul materialelor cerute de folosirea instrumentului sa nu fie prea mare.

In esenta, alegerea unui instrument va depinde de tipul de evaluare pe care-l dorim. In aceasta carte, cititorul va gasi instrumentele cele mai des folosite si criteriile de utilizare a lor pentru o evaluare confidenta a clientilor.

BIBLIOGRAFIE:

1. Dunn G. and Everitt B. Clinical Biostatistics. An Introduction to Evidence-Based Medicine, Edward Arnold, London, 1995
2. Froberg DG and Kane RL. Methodology for measuring health-state preferences - I: Measurement strategies, J. Clin. Epidemiol.1989,42:345-354
3. Salvador-Carulla L: Assessment instruments in psychiatry: description and psychometric properties, in Mental Health Outcome Measures, G Thornicroft & M Tansella (Eds.), Berlin: Springer, 1996
4. Heymans G, Wiersma E: Beitr•ge zur speziellen Psychologie auf Grund einder Massenuntersuchung. Zeitschrift f. Psychologie, 1909,51: 1-72
5. Wittenborn JR: A new procedure for evaluating mental hospital patients, J.Consult.Clin.Psychol.1950,14:500-501
6. Bobon D: Us et abuse des echelles d'evaluation en psychopathologie, Psychiatrie & Psychobiologie, 1987,2:379-385

7. Sartorius N: La clasification: une optique internationale, *Acta Psychiat. Belg.* 1983,83:88-103
8. Sartorius N: Making of a common language for psychiatry: development of the classification, behavioural and developmental disorders in the 10th revision of the ICD-10, *WPA Bulletin*, 1989, 1: 1
9. Popper K: *Objective Knowledge: An Evolutionary Approach*, Oxford: Clarendon Press, 1972
10. Dennis M, Ferguson B, Tyrer P: Rating instruments, in *Research Methods in Psychiatry*, C Freeman & P Tyrer (eds.), London: Gaskell, 1995
11. Bech P, Malt UF, Dencker SJ, Ahlfors UG et al: Scales for Assessment of Diagnosis and Severity of Mental Disorders, *Acta Psychiat.Scand.* 1993,87, Suppl. 372
12. Streiner DL, Norman GR: *Health Measurement Scales: A Practical Guide to Their Development and Use*, 2nd ed. Oxford: Oxford University Press, 1995
13. Hudelson PM: *Qualitative Research for Health Programmes*, Geneva: World Health Organization, 1996
14. Üstün B, Compton W, Mager D, Babor T, Baiyewu O, Chatterji S, Cottler L, Gogus A, Mavreas V, Peters L, Pull C, Saunders J, Smeets R, Stipek MR, Vradi R, Hasin D, Room R, van den Brink W, Regier D, Blaine J, Grant BF, Sartorius N: WHO study on the reliability and validity of the alcohol and drug use disorder instruments: overview of methods and results, *Drug and Alcohol Dependence*, 1997,47:161-169
15. Room R, Janca A, Bennett LA, Schmidt L, Sartorius N... Vradi R et al: WHO cross-cultural applicability research on diagnosis and assessment of substance use disorders: an overview of methods and selected results, *Addiction*, 1996,91: 199-220
16. Bech P: *Rating Scales for Psychopathology, Health Status and Quality of Life*, Berlin: Springer-Verlag, 1993
17. Zeally AK, Aitken RCB: Measurement of mood. *Proc.Roy.Soc.Med.* 1969,62:993-996
18. Endicott J, Spitzer RL, Fleiss JL, Cohen J: The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance, *Arch.Gen.Psychiatry*, 1976,33:766-771
19. Derogatis L, Cleary P: Confirmation of the dimensional structure of the SCL-90: A study in construct validation, *J.Clin.Psychol.* 1977,33:981-989
20. Osgood C, Suci G, Tannenbaum P: *The Measurement of Feeling*, University of Illinois Press, Urbana, 1957

21. Young RM, Knight RG: The Drinking Expectancy Questionnaire: A revised measure of alcohol related beliefs. *J.Psychopathol.Behav.Assess.* 1989,11:99-112;
22. Skinner HA, Allen BA: Alcohol dependence syndrome: Measurement and validation. *J.Abnorm.Psychol.*1982,91:199-209
23. Metz CE: Basic principles of ROC analysis, *Semin.Nucl.Med.* 1978,8: 283-298
24. Murphy JM, Berwick DM, Weinstein MC et al: Performance of screening and diagnostic tests, *Arch.Gen.Psychiatry*, 1987,44:550-555
25. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M: Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on early detection of persons with harmful alcohol consumption - II, *Addiction*, 1993,88:791-804
26. Bohn MJ, Babor TF, Kranzler HR: The Alcohol Use Disorders Identification Test (AUDIT): Validation of a screening instrument for use in medical settings, *J.Stud.Alcohol.* 1995, 56:423-432
27. Conigrave KM, Hall WD, Saunders JB: The AUDIT questionnaire: choosing a cut-off score, *Addiction*, 1995, 90:1349-1356
28. Dunham PJ: *Research Methods in Psychology*, New York: Harper & Row Publ. 1991
29. Zubin J: Use of research instruments in psychopathological assessment: Some historical perspectives, in *Measuring Mental Illness: Psychometric Assessment for Clinicians*, S.Wetzler (Ed.), Washington, DC: American Psychiatric Press, 1989
30. Bögels SM, van der Vleuten CPM, Blok G et al: Assessment and validation of diagnostic interviewing skills for the mental health professions, *J.Psychopath.Behav.Assessment*, 1995, 17:217-230
31. Reiser DE: The psychiatric interview, in *Review of General Psychiatry*, H.G. Goldman (Ed.), Cos Altos, CA: Lange Publ. 1984
32. Maier W, Philipp M, Buller R: The value of structured clinical interviews, letter to editor, *Arch.Gen.Psychiatry*, 1988, 45:963-964
33. Spitzer R, Fleiss JL, Burdock EI et al: The Mental Status Schedule: rationale, reliability, and validity, *Compr. Psychiatry*, 1964, 5:384-395
34. Endicott J, Spitzer RL: A diagnostic interview: the Schedule for Affective Disorders and Schizophrenia, *Arch.Gen.Psychiatry*, 1978, 35:837-844
35. Rubinson EP, Asnis GM: Use of structured interviews for diagnosis, in *Measuring Mental Illness: Psychometric Assessment for Clinicians*, S.Wetzler (Ed.), Washington, DC: American Psychiatric Press, 1989

36. Saghir MT: A comparison of some aspects of structured and unstructured interview, *Am.J.Psychiatry*, 1971, 128: 180-184.
37. World Health Organization: The Composite International Diagnostic Interview, core version 1.1. American Psychiatric Press, Washington, DC, 1993
38. World Health Organization: Schedules for Clinical Assessment in Neuropsychiatry (SCAN), American Psychiatric Press, Washington, DC, 1994
39. Pollock DC, Shanley DF, Byrne PN: Psychiatric interviewing and clinical skills, *Canadian J.Psychiat.* 1985, 30:64-68
40. Borus JF, Yager J: Ongoing evaluation in psychiatry: The first step toward quality, *Am.J.Psychiatry*, 1986, 143:1415-1419
41. Bögels SM, van der Vleuten CPM, Blok G et al: Assessment and validation of diagnostic interviewing skills for the mental health professions, *J.Psychopath.Behav.Assessment*, 1995, 17:217-230
42. Spitzer RL: Psychiatric diagnosis: Are clinicians still necessary? *Compr. Psychiatry*, 1983, 24:399-411
43. Griffin ML, Weiss RD, Mirin SM et al: The use of the Diagnostic Interview Schedule in drug-dependent patients, *Am.J Drug Alcohol Abuse*, 1987, 13:281-291
44. Pilkonis PA, Heape CL, Ruddy J, Serrao P: Validity in the diagnosis of personality disorders: The use of the LEAD standard, *Psychol.Assess.* 1991, 3:1-9
45. Kranzler HR, Kadden RM, Babor TF, Rounsaville BJ: Longitudinal, expert, all data procedure for psychiatric diagnosis in patients with psychoactive substance use disorders, *J.Nerv.Ment.Dis.*, 1994,182: 277-283
46. Landy FJ: Stamp collecting versus sciences, *American Psychologist*, 1986, 41:1183-1192
47. Young DW: Assessment of questions and questionnaires, *Meth.Inform.Med.* 1971,10:222-228
48. Allen JP, Columbus M: Assessing Alcohol Problems. A Guide for Clinicians and Researchers, NIAAA, Treatment Handbook Series 4, Bethesda, MD, 1995
49. Lyerly SB: Handbook of Psychiatric Rating Scales, National Institute of Mental Health, Rockville, MA, 1978
50. Cohen J: A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, 1960, 20:37-46
51. Fleiss JL: Statistical Methods for Rates and Proportions, New York: Wiley, 1981

52. Cottler LB, Compton WM, Brown L et al: The discrepancy interview protocol: A method for evaluating and interpreting discordant survey responses, *Int.J.Methods Psychiatr.Res.* 1994,4:173-182
 53. Grant BF, Hasin DS: *The Alcohol Use Disorders and Associated Disabilities Interview Schedule*, Rockville, National Institute on Alcohol and Alcoholism, 1991
 54. Chatterji S, Saunders JB, Vrasti R et al: Reliability of the alcohol and drug modules of the Alcohol Use Disorder and Associated Disabilities Interview Schedule-Aslcohol/Drug-Revised (AUDADIS-ADR): an international comparison, *Drug and Alcohol Dependence*, 1997,47:171-185
 55. Vrasti R, Grant BF, Chatterji S et al: Reliability of the Romanian version of the alcohol module of the WHO Alcohol Use Disorder and Associated Disabilities: Interview Schedule - Alcohol/Drug-Revised, *European Addiction Research*, 1998,4:144-149
 56. Marsella AJ, Westermeyer J: Cultural aspects of treatment: conceptual, methodological, and clinical issues and directions, in *Treatment of Mental Disorders: A Review of effectiveness*, N Sartorius et al (Eds.), Washington, DC: American Psychiatric Press Inc. 1993.
 57. Flaherty JA, Gaviria FM, Pathak D, et al: Developing instruments for cross-cultural psychiatric research, *J.Nerv.Ment.Dis.* 1988, 176:257-263
 58. Trotter RT: Ethnographic research methods for applied medical anthropology, in *Training Manual in Applied Medical Anthropology*, Carole Hill (Ed.), Washington, DC: American Anthropological Association, Special Publication No.27, 1991
 59. Marsella AJ, Kameoka V: Ethnocultural issues in the assessment of psychopathology, in *Measuring Mental Illness: Psychometric Assessment for Clinicians*, S Wetzler (Ed.), Washington, DC: American Psychiatric Association, 1989
 60. Friedman AS, Granick S: *Assessing Drug Abuse Among Adolescents and Adults: Standardized Instruments*, National Institute on Drug Abuse: Clinical Raport Series, Rockville, 1994
-