Partea I

ELEMENTE DE METODOLOGIE PEDAGOGICĂ A EVALUĂRII ŞI EXAMINĂRII

Metodologia evaluării, alături de metodologia predării-învăţării, face parte din sfera metodologiei pedagogice. Aceasta înseamnă că evaluarea nu este o componentă separată a procesului de învăţământ. Dimpotrivă, evaluarea este şi trebuie să fie privită ca o parte integrantă a procesului de învăţământ, care interacţionează cu obiectivele, cu conţinuturile şi, desigur, cu metodologia didactică. Necesitatea încorporării evaluării în ansamblul de componente şi acţiuni ce alcătuiesc procesul de învăţământ este evidenţiată, pe plan terminologic, prin folosirea sintagmei ***proces de predare, învăţare şi evaluare.***

Interdependenţa dintre metodele de predare şi metodele de evaluare constituie şi o cerinţă de ordin pedagogic, exprimată sintetic în principiul: ***evaluarea trebuie să reflecte predarea***, iar predarea, la rândul ei, trebuie să răspundă la feed-backul oferit de evaluare. Desigur, această interdependenţă nu trebuie să inducă o confuzie de planuri sau o ignorare a specificităţii fiecărei activităţi. Prin obiectivele şi prin funcţiile lor, metodele de evaluare nu rezultă dintr-o simplă preluare a metodelor de predare, ci dintr-o adaptare la cerinţele verificării şi aprecierii corecte şi semnificative a performanţelor studenţilor.

1. Metode de evaluare bazate pe verificarea orală

Metodele bazate pe verificarea (examinarea) orală fac parte din categoria mai generală a metodelor de comunicare verbală aplicate şi în procesul predării, aşa încât cerinţele generale privind realizarea comunicării profesor-studenţi se aplică şi cazul evaluării. În funcţie de tipul de comunicare şi de tipul de limbaj practicat (dialogat, monologat), se pot distinge două metode principale: conversaţia şi expunerea de evaluare.

1.1. Conversaţia de evaluare

Conversaţia de evaluare, numită de unii autori ***chestionarea orală***, constă într-o succesiune logică şi psihologică de întrebări şi răspunsuri, iniţiată şi coordonată de profesor, pe o temă dată. Conversaţia de evaluare poate fi folosită atât ca metodă de *evaluare continuă* (formativă) în cursul seminariilor, la care participă unul sau mai mulţi studenţi (inclusiv întreaga grupă), cât şi ca *metodă de examinare*, în sesiunile de examene, îndeosebi la examinările de tip colocviu.

1. **Cerinţe privind formularea întrebărilor**. În cadrul conversaţiei de evaluare, *întrebările* au rolul unor stimuli, al unor solicitări verbale menite să declanşeze procesele şi operaţiile intelectuale implicate în învăţarea cunoştinţelor şi a capacităţilor supuse verificării (evaluării). De aceea, problema centrală a conversaţiei de evaluare o constituie natura întrebărilor, modul de formulare şi de adresare (exprimare) şi, în legătură cu acestea, *tipul de solicitare* implicat în formularea verbală. O întrebare adresată memoriei studenţilor va genera un răspuns repetitiv,

reproductiv sau, după caz, un răspuns incomplet sau chiar nonrăspuns. Mai mult decât atât, un astfel de gen de întrebări va induce la studenţi un *stil de învăţare*, bazat pe memorare şi reproducere. Principalele cerinţe privind natura, formularea şi exprimarea întrebărilor pot fi sintetizate astfel:

a) Întrebările trebuie să fie formulate clar, într-un limbaj explicit, fără ambiguităţi de sens, iar solicitarea trebuie să fie sintetică şi directă. Când întrebarea se referă la o problemă mai complexă, sau când presupune precizări suplimentare care să direcţioneze gândirea studenţilor, se recomandă ca formularea întrebării să cuprindă două părţi: o parte explicativă, cu rol introductiv sau de lămurire a contextului şi o parte constând în întrebarea propriu-zisă, care trebuie să fie scurtă şi precisă.
b) Fiecare întrebare trebuie să exprime direct şi explicit natura şi conţinutul solicitării adresate studenţilor (studenţii să înţeleagă *ce li se cere*) şi genul de răspuns pretins (reproductiv sau creativ, scurt sau detaliat etc.). Cu alte cuvinte, *obiectul întrebării* trebuie să fie clar, iar, dacă apar dificultăţi, profesorul trebuie să verifice dacă studenţii au identificat şi au înţeles corect acest obiect.
c) Întrebările nu trebuie să sugereze răspunsul şi nici că un răspuns ar fi mai bun iar altul mai rău. Sugestia poate fi generată atât prin modul de formulare, cât şi prin modul de exprimare.
d) Întrebările nu trebuie să cuprindă simultan mai multe solicitări şi, mai ales, să nu cuprindă mai multe solicitări foarte diferite, care ar presupune zone diferite de concentrare mentală. În cazul în care verificarea urmăreşte capacitatea studenţilor de a stabili relaţii, de a efectua transferul de cunoştinţe, de a combina algoritmi diferiţi, se recomandă ca aceste aspecte să fie solicitate şi puse în evidenţă nu printr-o singură întrebare, ci printr-un lanţ de întrebări în care fiecare nouă întrebare să aibă în vedere şi răspunsurile date de studenţi la întrebările precedente.

2. **Cerinţe privind formularea şi acceptarea răspunsurilor**. Cealaltă componentă a conversaţiei de evaluare o constituie *răspunsurile date de studenţi*. De fapt, relaţia întrebare-răspuns formează o unitate indisociabilă; calitatea răspunsurilor date de studenţi nu poate fi apreciată separat de calitatea întrebărilor puse de profesor, iar fiecare nouă întrebare nu trebuie să fie formulată independent de răspunsurile date de studenţi la întrebările anterioare. Similar întrebărilor, şi răspunsurile studenţilor trebuie să respecte anumite cerinţe:

a) Răspunsurile trebuie să fie corecte şi complete, formulate fluent, coerent şi expresiv. Profesorul trebuie să descurajeze tendinţa studenţilor de a da răspunsuri "recitate", monotone, eliptice sau fragmentate.
b) Profesorul trebuie să lase studenţilor timp de gândire, de elaborare mentală a răspunsurilor, aşa-zisul dinamism al conversaţiei trebuie menţinut în limite acceptabile, fără a dezavantaja studenţii cu ritmuri intelectuale mai lente.
c) Răspunsurile greşite sau părţile greşite din răspunsuri trebuie corectate imediat şi, dacă este cazul, conversaţia de evaluare poate fi temporar întreruptă pentru explicarea de către profesor a naturii greşelilor, după care studentul poate fi solicitat să reformuleze răspunsul. În cazul în care frecvenţa

răspunsurilor incomplete, greşite sau a nonrăspunsurilor este mare, profesorul trebuie să adapteze nivelul de dificultate, să reformuleze întrebările,

În ceea ce priveşte modul de desfăşurare a conversaţiei de evaluare, opţiunea profesorului depinde de tipul de evaluare în care conversaţia este implicată. În cazul ***evaluărilor curente***, recomandarea generală este de a se folosi o conversaţie de tipul *comunicării de grup*, care implică participarea tuturor membrilor grupului. În acest caz conversaţia se practică frontal, cu întreaga grupă de studenţi şi presupune următoarele cerinţe:

- întrebările se adresează tuturor studenţilor din clasă, chiar dacă evaluarea vizează, în mod special, anumiţi studenţi;
- studenţii sunt antrenaţi în corectarea eventualelor erori ale colegilor, în completarea răspunsurilor sau în formularea altor variante de răspunsuri;

În cazul ***evaluărilor finale*** care folosesc *examinarea orală*, conversaţia de evaluare dobândeşte în mai mare măsură caracterul unei *chestionări orale*, la care participă un singur student. Este evident că, în acest caz, întrebările se adresează direct studentului şi ele sunt mult mai dependente de calitatea răspunsurilor precedente date de student.

Cu toate că valoarea formativă a conversaţiei de evaluare este incontestabilă, această metodă de evaluare prezintă şi unele dezavantaje. În primul rând, este o metodă care consumă mult timp, iar limitarea timpului poate conduce la o conversaţie superficială şi irelevantă, mai slabă decât verificările scrise, care au o putere mai mare de cuprindere sub aspectul numărului de studenţi evaluaţi. Ca dezavantaj, se menţionează şi caracterul mai puţin standardizat al evaluării prin conversaţie, dificultatea aplicării unor bareme riguroase şi, prin urmare, riscul unui subiectivism mai accentuat al profesorului.

2. Expunerea liberă a studentului

Expunerea liberă a studentului face parte din aceeaşi categorie a metodelor bazate pe comunicarea verbală orală şi se supune regulilor generale ale acestui tip de comunicare. Spre deosebire de conversaţie, prestaţia studentului nu mai apare ca o suită de răspunsuri la o succesiune de întrebări, ci ca o *prezentare compactă*, în limbaj oral, a unei teme mai ample, care se pretează la o expunere unitară şi coerentă. Opţiunea pentru această metodă de evaluare depinde de două repere principale:

a) ***obiectivele*** urmărite în evaluare, metoda fiind recomandată atunci când obiectivele vizează evaluarea capacităţii studentului de construcţie a unui mesaj coerent, de combinare şi sistematizare a cunoştinţelor, de exprimare fluentă şi expresivă, de înlănţuire logică a argumentaţiei;
b) ***conţinuturile*** care se evaluează, metoda fiind recomandată în cazul conţinuturilor cu o structură logică şi psihologică unitară, care nu pot fi fragmentate fără a se pierde o parte semnificativă din logica lor internă sau care, prezentate într-o formă scurtă, nu reflectă nivelul real de însuşire a lor de către studenţi.

Foarte important pentru valoarea metodei este modul în care profesorul formulează subiectul (tema) expunerii. Se recomandă ca subiectul sau tema să aibă un caracter problematizat, să implice stabilirea de relaţii, elaborarea de sinteze şi sistematizări, să nu depindă excesiv de informaţii memorizate şi ale căror lacune ar

putea să fragmenteze expunerea sau chiar să o blocheze. Profesorul trebuie să încurajeze continuitatea expunerii, să nu întrerupă studentul decât în cazul unor erori importante sau al devierii expunerii într-o direcţie greşită. În funcţie de contextul evaluării, expunerea studentului poate fi sprijinită pe schiţe scrise, în care studentul îşi structurează cele mai importante puncte de reper, cum este cazul examinărilor orale.

În aplicarea acestei metode, profesorul trebuie să aprecieze (şi să încurajeze) nu numai corectitudinea ştiinţifică a prestaţiei studentului, dar şi aspectele referitoare la continuitatea, coerenţa şi fluenţa expunerii, expresivitatea, capacitatea studentului de a valorifica maximal cunoştinţele pe care le deţine, de a evita "căile închise", de a pune în valoare ceea ce ştie şi de a nu pătrunde în zone în care are lacune.

2. Metode de evaluare bazate pe verificarea scrisă

Metodele de evaluare bazate pe verificarea scrisă rămân în sfera metodelor de comunicare verbală, dar asumând rigorile specifice acestui tip de comunicare. Principalul avantaj al acestor metode - avantaj ce explică larga lor utilizare - constă în posibilitatea verificării şi evaluării simultane, într-un timp relativ scurt, a unui număr mare de studenţi. La aceasta se adaugă şi posibilitatea unei evaluări cu un caracter mai riguros şi mai obiectiv.

2.1. Lucrările scrise de verificare curentă

Lucrările scrise de verificare curentă, sunt tehnici de verificare şi evaluare folosite pe parcursul procesului de învăţământ, integrate de obicei în seminarii şi activităţi practice, având rolul unui sondaj frontal prin care se verifică nivelul şi calitatea pregătirii curente a studenţilor. Acest tip de lucrări au un rol important în cadrul *evaluării formative* (continue), întrucât ele servesc nu numai la evaluarea şi notarea studenţilor, dar şi ca tehnici de cunoaştere a nivelului şi calităţii cunoştinţelor asimilate de studenţi în cadrul unui capitol sau al unei teme complexe, ca bază pentru abordarea unui nou capitol al programei.

2.2. Lucrările scrise de evaluare finală

Lucrările scrise de evaluare finală sunt tehnici evaluare care se aplică la sfârşitul unei perioade compacte a procesului de învăţământ, de regulă semestrul, având rolul de a pune în evidenţă nivelul şi calitatea pregătirii studenţilor pe ansamblul perioadei respective. Este important de subliniat că prin aceste lucrări se verifică şi se evaluează nu suma cunoştinţelor studenţilor, ci capacitatea de a opera cu cunoştinţele învăţate, de a stabili relaţii şi interdependenţe, de a rezolva probleme şi de a realiza aplicaţii.

Privite în ansamblul lor, tehnicile de verificare scrisă prezintă o serie de avantaje care explică, de altfel, larga lor utilizare. Avantajul principal constă în posibilitatea evaluării şi notării unui număr mare de studenţi. Pe lângă acest avantaj de natură cantitativă, tehnicile de verificare scrisă au şi avantaje de factură pedagogică. Ele favorizează activitatea independentă a studenţilor, le permite un ritm de lucru mai apropiat de posibilităţile fiecăruia, diminuează emoţiile specifice examinării orale. De asemenea, lucrările scrise se pretează unei evaluări mai obiective, bazată pe analiza comparativă postevaluare, pe folosirea unor bareme sau standarde de corectare.

Pe de altă parte, tehnicile de verificare scrisă au şi dezavantaje. Principalul dezavantaj aparţine însăşi comunicării scrise, care nu permite un feed-back direct şi operaţional. Actul emiterii (al elaborării lucrărilor de către studenţi) şi actul receptării (al corectării lucrărilor de către profesor) sunt separate în timp şi spaţiu. Studentul nu poate reveni asupra celor scrise, nu poate aduce explicaţii suplimentare, nu beneficiază de întrebări ajutătoare şi nici de confirmarea sau infirmarea directă sau indirectă a corectitudinii răspunsurilor date. La rândul lui, profesorul nu poate solicita precizări suplimentare din partea studentului şi nu poate corecta operativ erorile. De aceea, o cerinţă aproape obligatorie este ca rezultatele lucrărilor scrise să fie prezentate studenţilor, discutate cu studenţii, cu precizarea principalelor calităţi, dar şi a erorilor tipice, a rezolvărilor greşite şi a concluziilor privind pregătirea viitoare.

2.3. Evaluarea prin tehnica testelor

Utilizarea tehnicii testelor în verificarea şi evaluarea nivelului de pregătire a studenţilor este una dintre soluţiile tot mai larg aplicate de creştere a caracterului obiectiv şi semnificativ al evaluării şcolare şi de asigurare a unei comparabilităţi sporite a evaluărilor efectuate pe eşantioane de studenţi diferite, de către evaluatori diferiţi şi în perioade diferite. Testele sunt recomandate ca o tehnică de diminuare a erorilor subiective de evaluare, dar şi ca o tehnică prin care se pot pune în evidenţă tipurile de dificultăţi şi de erori, fiind astfel o bună tehnică de evaluare formativă.

În general, testul este o ***probă standardizată*** prin care se determină nivelul şi caracteristicile unor procese şi însuşiri psihice la o anumită populaţie (eşantion), precum şi poziţia relativă a unei persoane în cadrul populaţiei respective. Testul este o probă şi în sensul că "pune la încercare", solicită subiectului "să probeze" o anumită capacitate, o anumită performanţă sau o anumită însuşire psihică.

Testul implică o *examinare identică* pentru toţi subiecţii şi un mod de evaluare şi de notare de asemenea identic. Prin aceasta, testul este sau intenţionează să fie un ***instrument de măsură***, care foloseşte un anumit *etalon*, în baza căruia se poate stabili nu numai nivelul absolut al performanţei testate, dar şi nivelul sau poziţia unei persoane în cadrul populaţiei din care face parte, din punctul de vedere al performanţei respective. Este de precizat că elaborarea şi etalonarea testelor implică demersuri şi tehnici destul de complexe, care nu se află întotdeauna la îndemâna profesorului. De regulă, testele standardizate şi etalonate se construiesc de către echipe de specialişti, în instituţii specializate. De aceea, în procesul de învăţământ şi în activitatea curentă de evaluare realizată de profesor la clasă se poate vorbi mai curând de folosirea ***tehnicii testelor***, decât a unor teste psihologice propriu-zise.

Denumirea cea mai generală a testelor folosite în evaluarea educaţională este aceea de ***test educaţional***. În practică însă, testele nu măsoară sau nu pot măsura tot complexul de rezultate ale procesului educaţional, cel mai frecvent (şi mai accesibil) fiind *domeniul cognitiv*, aşa încât se preferă, de cele mai multe ori, termenul de ***test de cunoştinţe***. De asemenea, spre a se sublinia rolul testului ca tehnică examinare şi de măsurare a performanţelor şcolare, se foloseşte şi termenul de ***test docimologic***. Dacă se consideră că rigorile privind etalonarea, validitatea etc. nu sunt îndeplinite, atunci este preferabil termenul de ***probă de cunoştinţe*** sau ***probă docimologică.***

Sub aspect constitutiv, unitatea de bază a testului este ***itemul***. Acesta este o

solicitare punctuală (o întrebare, o problemă), precis delimitată, pe care subiectul trebuie să o îndeplinească. Fiecare item corespunde unui criteriu (tip de capacitate sau de performanţă), aşa încât itemii nu sunt entităţi izolate, ci sunt organizaţi în *seturi sau baterii de itemi,* în interiorul cărora itemii sunt ierarhizaţi pe tipuri de solicitări şi niveluri de dificultate.

Din punctul de vedere al *tipului de răspunsuri solicitate*, itemii pot fi de două feluri: itemi cu ***răspunsuri deschise***, la care răspunsurile trebuie elaborate de studenţi şi itemi cu ***răspunsuri închise***, la care studenţii trebuie să aleagă răspunsurile corecte dintr-o listă de variante de răspunsuri preformulate.

1. **Itemii cu răspunsuri deschise** pot fi construiţi prin două tehnici principale: ***tehnica răspunsului scurt***, în care răspunsul corect şi complet are o formulare scurtă şi precisă, alcătuită din date, nume, termeni, formule, factori etc., care pot fi enunţate (scrise) ca atare, fără completări sau precizări suplimentare; ***tehnica răspunsului elaborat***, în care răspunsul corect şi complet presupune o construcţie verbală cel puţin de nivelul unei fraze (în sens gramatical), care implică folosirea mai multor termeni, noţiuni, date etc. În funcţie de natura solicitărilor, răspunsurile pot să fie de ***tip algoritmic***, când există o structură cunoscută de reguli şi operaţii ce trebuie aplicată de student (de exemplu, definiţiile, clasificările, procedurile, formulele etc.), sau de ***tip euristic***, când se cer studenţilor interpretări, exemplificări, aprecieri şi, în general, sarcini pentru care nu există o soluţie unică şi certă care să fi fost anterior învăţată.

2. **Itemii cu răspunsuri închise** cuprind două sau mai multe variante de răspunsuri, dintre care una sau mai multe sunt corecte (adevărate), iar celelalte greşite (false). Desigur, variantele greşite nu trebuie să conţină absurdităţi, aberaţii sau erori "grosolane", uşor identificabile, ci soluţii plauzibile, adevăruri aparente care, dacă sunt alese de studenţi, traduc erori tipice, lacune cognitive, imprecizii, confuzii, generalizări pripite etc. Cu alte cuvinte, varianta greşită trebuie "să pară corectă", să conţină o anumită "capcană" care să surprindă fie învăţarea superficială, fie o eroare de gândire, de înţelegere sau de aplicare a cunoştinţelor. De aceea, unii autori fac deosebire între *răspunsuri,* care sunt variantele corecte, şi *distractori*, care sunt variantele greşite, aceştia din urmă având rolul de a distrage atenţia studenţilor de la răspunsurile corecte în cazul în care calitatea slabă a învăţării nu permite identificarea certă a răspunsului corect.

Există o mare varietate de tehnici de construire a itemilor cu răspunsuri închise. În funcţie de numărul variantelor de răspuns, există itemi cu *alegere dihotomică sau dualăşi* itemi cu *alegere multiplă*.

a) ***Itemul cu alegere dualăconstă*** dintr-un enunţ (o afirmaţie sau o negaţie, o ipoteză, o regulă etc.), iar studentul este solicitat să aleagă între două variante de tipul: corect-greşit, adevărat-fals, acord-dezacord, pozitiv-negativ, da-nu etc.

b) ***Itemul cu alegere multiplăconstă*** dintr-o întrebare, o problemă sau un enunţ pentru care se oferă mai multe (cel puţin 3) variante de răspunsuri. Aceste variante pot să aibă forma unor *răspunsuri complete preformulate* sau, în cazul problemelor, a rezultatului final al rezolvării, sau forma unor *scale de evaluare cantitativă* (foarte mult, mult, puţin etc.), *calitativă* (foarte bine, bine, mai puţin bine etc.) sau *temporală* (întotdeauna, deseori, uneori etc.). În privinţa numărului şi naturii variantelor de răspunsuri corecte/greşite, există mai multe opinii, fiecare evidenţiind avantaje şi

dezavantaje. Tehnica alegerii multiple prezintă două varietăţi principale: *alegerea răspunsului corect*, caz în care în care există o singură variantă corectă şi *alegerea celui mai bun răspuns*, caz în care mai multe sau toate variantele sunt "aproximativ corecte", mai mult sau mai puţin discutabile, dar una singură are cel mai ridicat "grad de adecvare" cu premisa enunţată în item.

Opţiunea pentru o singură variantă corectă prezintă riscul "ghicirii" sau pur şi simplu al alegerii din întâmplare a variantei corecte, mai ales când numărul variantelor este redus. De pildă, la 3 variante, există o probabilitate de 30% ca răspunsul corect să fie ales la întâmplare, dar probabilitatea "ghicirii" poate fi mult mai mare, dacă avem în vedere că majoritatea studenţilor ştiu "câte ceva" despre conţinuturile testului.

2.4. Chestionarul de evaluare

Chestionarul de evaluare constă într-o succesiune logică şi psihologică de întrebări şi răspunsuri prin care se urmăreşte verificarea şi evaluarea nivelului şi calităţii achiziţiilor studenţilor pe o gamă largă de obiective şi conţinuturi. Spre deosebire de testul de cunoştinţe, chestionarul de evaluare nu este o probă standardizată, rolul lui este de *a colecta informaţii* şi nu de a măsura (în sensul în care se face măsurarea prin testele de cunoştinţe).

În construirea chestionarului, mai importante decât standardizarea şi tipizarea itemilor sunt *fluenţa şi coerenţa internă a succesiunii de întrebări, legătura logică şi de conţinut între o întrebare şi alta*. Chiar dacă se prezintă ca un formular scris, chestionarul păstrează în bună măsură caracterul unui dialog sau al unui discurs în care întrebările nu sunt entităţi izolate, ci fac parte dintr-un demers cognitiv şi comunicativ coerent, cu o anumită continuitate de conţinut şi înlănţuire logică. De aceea, chestionarul de evaluare se elaborează în jurul unei anumite teme, la încheierea unui anumit capitol din programă, când sunt necesare şi posibile sinteze, transferuri de cunoştinţe, comparaţii, generalizări. Chestionarul poate fi şi un bun *instrument de autoevaluare* pentru studenţi. In acest scop, profesorul poate da studenţilor chestionare ca ghid în pregătirea seminariilor, lucrărilor practice sau a examenelor.

3. Metode de evaluare bazate pe verificarea practică

3.1. Aplicaţiile

Aplicaţiile sunt metode de evaluare care solicită studenţilor *utilizarea practică a cunoştinţelor* în efectuarea unor lucrări practice care presupun exersarea cunoştinţelor respective. Evident, aplicaţiile se pretează la acele discipline şi conţinuturi pentru care importantă nu este cunoaşterea lor teoretică, ci folosirea în activităţi practice şi, mai ales, în activităţi profesionale.

Trebuie subliniat că nu orice utilizare practică a cunoştinţelor are caracterul unei aplicaţii. De exemplu, la matematică, se poate cere studenţilor să aplice una sau mai multe formule în rezolvarea unei probleme, ceea ce nu reprezintă însă o aplicaţie ca metodă de evaluare. Pentru a fi o aplicaţie, cunoştinţele de matematică trebuie utilizate într-o activitate concretă, de regulă cu caracter de *activitate profesională*. De exemplu, la statistică socială, studenţilor de la sociologie li se poate cere folosirea instrumentului statistic în analiza şi interpretarea datelor unei investigaţii sociologice

concrete. Rezultă că aplicaţiile presupun exersarea cunoştinţelor în *situaţii profesionale reale*, fiind nemijlocit legate de tipul de activitate didactică bazat pe lucrări practice.

3.2. Analiza produselor activităţii studenţilor

Activitatea de învăţare desfăşurată de studenţi se materializează deseori în produse, în lucrări, în obiecte fizice care pot constitui un bun reper pentru verificarea şi evaluarea cunoştinţelor, capacităţilor şi deprinderilor dobândite în procesul de învăţământ. Produsul activităţii are avantajul că sintetizează foarte bine un complex de caracteristici incluzând domeniul *cognitiv* (cunoştinţe, capacităţi), domeniul *motivaţional-atitudinal* (motivaţii, interese, atitudini) şi domeniul *psiho-motor, de aplicare şi execuţie* (deprinderi, abilităţi).

În sens larg, prin produs se înţelege orice rezultat fizic al activităţii studenţilor realizat de ei în cadrul sau în legătură cu procesul de învăţământ. În acest sens, sunt produse ale activităţii: lucrările scrise, referatele, lucrările de laborator sau atelier etc. Desigur însă că analiza produselor activităţii este o *metodă specifică de evaluare* în cazul tipurilor de activitate didactică (lucrări practice, laboratoare, seminariile cu caracter aplicativ) ce presupun prin obiectivele şi conţinutul lor realizarea de produse. În acest caz, realizarea produsului reprezintă principala modalitate de învăţare, iar caracteristicile produsului principalul criteriu de evaluare.

Pentru a oferi o evaluare semnificativă, analiza produselor activităţii studenţilor trebuie să se întemeieze pe repere şi criterii clare şi pertinente. Desigur, acestea sunt în bună măsură dependente de natura produsului şi a activităţii didactice în care produsul a fost realizat. Dintre criteriile de evaluare cu o aplicabilitate mai generală se pot menţiona:

- gradul de corespondenţă cu obiectivele sau cu parametrii proiectaţi, în baza cărora produsul a fost realizat;
- aspectele tehnice sau procedurale ale realizării produsului: aplicarea tehnicilor şi procedeelor recomandate, calitatea operaţiilor efectuate;
- aspectele estetice ale produsului;
- aspectele relevante pentru atitudinea studentului în procesul executării produsului, cu referire la acurateţea execuţiei, atenţia acordată detaliilor aparente, temeinicia realizării, exigenţa în autocontrolul calităţii etc.

Partea a II-a

ELEMENTE DE TEORIA ŞI PRACTICA NOTĂRII ŞCOLARE

Notarea şcolară este ca actul prin care se atribuie note sau calificative prestaţiilor studenţilor în urma diferitelor forme de evaluare. Notarea exprimă şi oficializează nivelul performanţelor atinse de studenţi în diferitele etape ale procesului de învăţământ. Importanţa notării şcolare nu se limitează însă la cadrul strict pedagogic al procesului de învăţământ. Notarea este nu numai un act pedagogic, dar şi unul cu semnificaţii şi consecinţe care depăşesc graniţele instituţiei de învăţământ şi

influenţează statutul social şi profesional al studenţilor.

Calitatea notării îşi pune amprenta pe calitatea întregului act de evaluare, aşa încât sistemul de evaluare în ansamblul său nu poate fi ameliorat şi perfecţionat dacă nu se asigură un sistem de notare corect. Din această perspectivă, prezentăm în continuare o sinteză a principalelor aspecte privind notarea şcolară, valoarea şi limitele acesteia, condiţiile aplicării corecte a sistemului de notare.

1. Notarea şcolară ca act de măsurare

Dacă luăm în considerare accepţiunea generală a conceptului de măsurare - ca acţiune de determinare a valorilor unei mărimi şi de exprimare a acestor valori printr-o serie numerică - atunci *notarea şcolară poate fi tratată ca un act de măsurare.* Fireşte, nu vom putea asimila notarea şcolară cu măsurarea caracteristicilor fizice sau cu măsurarea formală din matematică. Ea este însă destul de apropiată de măsurarea din psihologie sau din sociologie, iar adoptarea unor concepte, condiţii şi tehnici din aceste domenii se poate dovedi utilă pentru ameliorarea notării şcolare.

1.1. Particularităţi

Pentru a evita eventualele confuzii între actul măsurării şi cel al evaluării prin note şcolare, este necesar să subliniem că:

1. Notarea şcolară nu este *stricto - senso* un act de măsurare întrucât notele şcolare, deşi exprimate numeric, nu reflectă mărimi absolute ci *ierarhii*, poziţii sau ranguri, aşa încât *între mărimile caracteristicii măsurate şi seria numerică folosită nu există o relaţie de identitate izomorfă,* în sensul că în timp ce numerele cresc sau descresc cu intervale egale, distanţele reale care separă o notă şcolară de alta nu sunt egale pe diferitele paliere ale scării de notare (de pildă, intervalul dintre notele 4 şi 5 nu este egal cu intervalul dintre notele 9 şi 10, iar acesta din urmă nu este egal cu intervalul dintre 5 şi 6 ş.a.m.d.);

2. Exprimând *relaţii de ordine* şi nu cantităţi de sine stătătoare, *notele şcolare nu au semnificaţia unor valori absolute* şi nici nu poate fi stabilit un punct sau nivel *zero absolut* (sau natural) al caracteristicii măsurate - care să desemneze absenţa caracteristicii sau inexistenţa unui nivel mai mic decât cea mai mică valoare a caracteristicii măsurate - de unde rezultă două consecinţe deosebit de importante:

a) relaţia dintre numerele care exprimă notele şcolare *nu este aditivă*, ceea ce înseamnă că numărul care exprimă două mărimi reunite nu este egal cu suma a două numere care exprimă fiecare din cele două mărimi. Foarte simplu spus, aceasta înseamnă că doi elevi de nota 5 nu sunt egali cu un elev de nota 10, sau că o performanţă de nota 9 nu este suma unor performanţe distincte cotate cu 2, 6 şi 1 (deşi 5+5=10, iar 9=2+6+1).
b) notele şcolare nu exprimă nici produse şi nici raporturi între mărimile pe care le desemnează; astfel, o performanţă de nota 8 nu poate fi tratată ca de două ori mai mare decât una de 4, iar o performanţă de nota 3 nu exprimă a treia parte dintr-o performanţă de nota 9.

3. Întrucât acordarea notelor implică un act de raportare la anumite criterii şi la anumite valori, relaţia dintre obiectul evaluării şi sistemul de măsurare nu este directă,

ci *mijlocită de semnificaţii*, aşa încât nota acordată depinde atât de proprietăţile obiectului evaluării, cât şi de sensul pe care evaluatorul sau sistemul de evaluare îl acordă proprietăţilor respective. Aceasta deoarece nota şcolară măsoară nu doar cantităţi, ci şi *calităţi*, ea exprimă nu numai ordine de mărime (intensitate, frecvenţă, volum), ci şi *atribute* cu semnificaţia de pozitiv sau negativ, favorabil sau nefavorabil.

Menţinând analiza în limitele acestor particularităţi ale notării şcolare, vom putea în continuare să supunem atenţiei posibilităţile de control al erorilor de notare ce rezultă din *considerarea notării şcolare ca un act de măsurare*.

1.2. Structura actului de notare

În teoria generală a măsurării se consideră că structura actului de măsurare cuprinde trei componente: *obiectul* de măsurat, *etalonul* de măsură şi *regulile de atribuire a valorilor*. Aceeaşi structură poate fi identificată şi în notarea şcolară, cu unele particularităţi pe care le vom enunţa în cele ce urmează. Analiza structurii actului de notare este operaţia prin care se identifică, se defineşte şi se controlează fiecare element al structurii, inclusiv sub aspectul posibilităţilor de apariţie a erorilor.

Trebuie spus, în acest context, că o mare parte a erorilor de apreciere îşi au originea tocmai în slaba definire şi corelare a componentelor actului de evaluare. În planul practic al evaluării şcolare, analiza structurii actului de notare presupune ca profesorul să răspundă următoarelor întrebări:

- *Ce anume evaluează*? ceea ce înseamnă a defini *obiectul* evaluării (notării);
- *Ce sistem de notare foloseşte*? ceea ce înseamnă a preciza caracterul şi natura *etalonului* folosit;
- *Pe baza căror reguli atribuie notele şcolare*? ceea ce înseamnă a stabili ce relaţii există între caracteristicile obiectului evaluării şi notele acordate.

Răspunsul la oricare dintre aceste întrebări ar presupune incursiuni în problematica destul de complexă a măsurării şi evaluării în ştiinţele socio-umane. Ne vom limita însă la aspectele nemijlocit legate de problema evaluării şcolare.

1. ***A defini obiectul evaluării*** înseamnă a stabili *care sunt capacităţile, performanţele, atitudinile avute în vedere şi ce caracteristici ale acestora sunt transpuse în note şcolare*. Trebuie subliniat că în cazul evaluării şcolare obiectul evaluării este unul complex, cu numeroase componente calitative care, în planul măsurării, sunt rareori accesibile în mod direct. De cele mai multe ori, evaluarea şcolară are caracterul unei *măsurări indirecte*, ceea ce înseamnă că se notează nu caracteristicile ca atare, ci *manifestările observabile ale acestora* sau care pot fi abordate printr-un instrument sau printr-o metodă de investigaţie.

Rezultă că a defini obiectul evaluării înseamnă a stabili indicatorii empirici, care vor fi înregistraţi şi notaţi. În tehnicile moderne de evaluare, aceşti indicatori sunt numiţi ***descriptori de performanţă***, iar formularea lor este similară modului de formulare a obiectivelor operaţionale, adică precizându-se comportamentul şi performanţa care vor fi înregistrate şi măsurate. Rezumând, vom spune că definirea obiectului evaluării are două componente:

a) capacitatea sau subcapacitatea care este supusă evaluării şi
b) indicatorii sau descriptorii prin care capacitatea este pusă în evidenţă.

Pe de altă parte, definirea obiectului evaluării trebuie să fie făcută de manieră să

permită *sesizarea variaţiilor* cantitative şi calitative ale caracteristicilor evaluate. Sub acest aspect, profesorul trebuie să găsească soluţii la două probleme:

- *problema nivelului minim şi maxim de dificultate,* adică a amplitudinii intervalului de evaluare, aşa încât acesta să corespundă cât mai bine scării de notare folosită;
- *problema metodei sau tehnicii potrivite* cu tipul de indicatori evaluaţi. Ca regulă generală se recomandă combinarea mai multor modalităţi de verificare şi evaluare (orală, scrisă, practică ş.a.m.d.).

2. Al doilea element al analizei structurii actului de notare este ***natura sistemului de notare*** *şi a* ***etalonului*** *folosit în evaluare*. După cum se cunoaşte, etalonul este un *fragment standardizat al mărimii ce se măsoară* care serveşte drept numitor comun pentru toate măsurările efectuate pe aceeaşi dimensiune. Se ridică întrebarea: Îndeplineşte scala de notare şcolară condiţiile unui etalon de măsură? Răspunsul cel mai potrivit nu este în termeni de DA sau NU, ci în termeni de cât şi în ce condiţii notarea poate să devină un etalon de măsură.

În acest sens, trebuie subliniat că scala de notare are caracterul unui *etalon convenţional* (şi nu natural) ceea ce înseamnă că relaţia dintre obiectul evaluării şi sistemul de măsură este indirectă, depinzând de calitatea metodelor sau instrumentelor prin care se face evaluarea. Dacă vom compara scala de notare cu modalităţile de etalonare utilizate în psihologie, atunci vom constata că practic ea apare ca o *scală bazată pe cote normalizate*, în sensul că, de regulă, profesorii acordă notele în ponderi care aproximează o distribuţie normală.

3. Al treilea element al structurii actului de notare îl constituie ***regulile de atribuire a notelor*** sau, cu alte cuvinte, ***criteriile*** pe baza cărora o anumită performanţă este pusă în legătură cu un anumit nivel al scării. Prin regulile de acordare a notelor, profesorul trebuie să asigure de fapt o corespondenţă cât mai deplină între nivelurile de performanţă - identificate prin tehnicile de verificare - şi sistemul de note (ca sistem numeric).

Problema este *la ce nivel de măsurare se situează evaluarea şcolară* şi deci ce operaţii sunt permise cu notele şcolare acordate. Deseori se neglijează acest aspect, ceea ce conduce la scări diferite şi necomparabile şi la calcule nepermise. În acest sens trebuie subliniat că, prin cele mai multe caracteristici ale sale, notarea şcolară nu depăşeşte ***nivelul ordinal al măsurării***. Ce consecinţe are acest fapt?

a) În primul rând, aceea că notele şcolare nu reflectă mărimi absolute ci *ranguri*, poziţii sau ierarhii pe o scară ordonată crescător sau descrescător, indiferent că se folosesc numere sau calificative. Aceasta înseamnă că o anumită notă nu are semnificaţie decât raportată la alta mai mică sau mai mare.
b) În al doilea rând, în sistemul de notare *nu poate fi stabilit (nu are sens) un punct zero natural* care să aibă semnificaţia de absenţă a caracteristicii evaluate sau de nivel dincolo de care nu există o valoare mai mică. De aceea, scala de notare nu include nivelul zero. De aici rezultă o consecinţă deosebit de importantă: aceea *că notele nu pot fi însumate, înmulţite sau împărţite în acelaşi mod în care aceste operaţii se fac cu cantităţile fizice*. Este evident că două performanţe distincte reunite nu conduc la o performanţă care să poată fi cotată cu dublul notelor acordate separat (două performanţe de 5 nu formează

o performanţă de 10). Tot astfel, un student de nota 8 este cu 4 unităţi superior unuia cu nota 4 şi nu de două ori mai bun.

c) În al treilea rând, în timp ce sistemul numerelor naturale are intervale egale între numere pe toată întinderea lui, intervalele reale dintre notele şcolare nu sunt egale pe diferitele niveluri ale scării de notare.

Toate aceste caracteristici impun anumite restricţii privind operaţiile ce se pot efectua cu notele şcolare. Astfel, efectuarea unor operaţii de adunare, împărţire sau înmulţire cu notele trebuie privită cu circumspecţie şi are întotdeauna o anumită doză de arbitrar. De exemplu, o medie şcolară de 6 ce rezultă din notele 3 şi 9 este în mod evident nesemnificativă şi induce în eroare evaluarea studentului. Într-o astfel de situaţie ar fi necesară cel puţin încă o evaluare pentru ca media calculată să fie minimal reprezentativă.

2. Condiţiile de corectitudine a notării

Soluţii utile şi eficiente de control al erorilor subiective de evaluare le oferă analiza condiţiilor pe care trebuie să le îndeplinească o măsurare corectă, condiţii studiate mai ales în psihologie, cu referire la elaborarea testelor psihologice, precum şi în managementul resurselor umane cu referire la evaluarea performanţelor profesionale. Considerăm că abordarea evaluării şcolare şi în special a notării prin prisma acestor condiţii merită mai multă atenţie şi poate oferi un ghid valoros al corectitudinii în evaluarea şcolară.

2.1. Validitatea notării

În general, *validitatea* este proprietatea unui sistem de măsurare *de a măsura ceea ce şi-a propus şi nu altceva*. Deşi validitatea pare o condiţie de la sine înţeleasă, nu este deloc aşa, iar cele mai multe erori sunt erori de validitate. Se poate spune că pentru evaluarea şcolară asigurarea validităţii sistemului de notare, inclusiv la nivelul metodelor şi instrumentelor de verificare/evaluare, este o problemă centrală a corectitudinii evaluării. Un sistem de notare este nevalid când ceea ce măsoară este *altceva* decât ceea ce urmăreşte să măsoare. Este cazul de pildă când urmărim să apreciem prin note învăţarea unor capacităţi complexe (analiză, sinteză, explicare, interpretare), dar performanţele pe baza cărora facem notarea sunt rezultatul altor mecanisme sau procese de învăţare, de pildă memoria, reproducerea.

Din punctul de vedere al problemelor pe care urmăreşte să le rezolve, evaluarea validităţii se efectuează pe trei tipuri de validitate: de conţinut, conceptuală şi predictivă.

1. ***Validitatea de conţinut*** se referă la gradul în care indicatorii sau descriptorii de performanţă folosiţi în evaluare alcătuiesc o *selecţie reprezentativă* din ansamblul indicatorilor ce acoperă caracteristica evaluată. Nevaliditatea este generată în acest caz de doi factori:

- indicatorii aleşi nu sunt pertinenţi pentru nivelurile caracteristicii evaluate;
- numărul indicatorilor este prea mic.

Asigurarea validităţii de conţinut presupune în principal intervenţii asupra instrumentelor de evaluare, ea se asigură în primul rând în etapa de verificare. De reţinut este faptul că verificările sumare, cu un număr şi o diversitate redusă de

probleme, puse în condiţii omogene conduce aproape sigur la o evaluare nevalidă.

2. ***Validitatea conceptuală*** numită şi *de construct* depinde de calitatea conceptelor pe care se întemeiază actul evaluării şi în primul rând de claritatea cu care a fost definit obiectul evaluat. Astfel, când capacitatea sau performanţa care se evaluează este precis definită, prin indicatori operaţionali, atunci ea va putea fi evaluată în acelaşi sens şi în acelaşi mod fie prin tehnici diferite, fie de către evaluatori diferiţi. În acest fel vor deveni posibile *comparaţiile* între evaluări diferite şi verificarea fiecărei evaluări prin raportări la celelalte.

Această tehnică de validare prin comparare este cunoscută şi sub denumirea de ***validitate concurentă*** şi este larg folosită în sistemele de moderare interevaluatori. Astfel, profesorul poate să şi evalueze validitatea notării, oferind altui profesor conceptul şi indicatorii pe care i-a folosit, dar nu şi instrumentele utilizate şi să compare în final rezultatele proprii cu ale celuilalt profesor.

3. ***Validitatea predictivă*** este probabil cea mai importantă formă de validitate pentru evaluarea şcolară întrucât ea desemnează capacitatea sistemului de notare de a prognostica evoluţia viitoare a studentului din punctul de vedere al caracteristicilor evaluate. Ea se mai numeşte şi *validitate de criteriu* sau ***criterială*** deoarece se determină prin raportare la un *criteriu exterior* în raport cu cel pe baza căruia s-a făcut evaluarea. Acest criteriu este de regulă o activitate, o performanţă sau o competenţă care reprezintă un mod de manifestare reală a caracteristicii evaluate.

Pentru evaluarea şcolară acest criteriu este şi trebuie să fie în principal *succesul/insuccesul studentului în acele activităţi care reprezintă o practicare, o exersare a capacităţilor pentru care a fost notat.* Astfel, o notă şcolară este validă, în sens predictiv, dacă este confirmată de rezultatele obţinute de student în afara procesului de învăţământ, la un examen, la un concurs, în activitatea profesională. Validitatea predictivă poate fi evaluată prin folosirea aşa-numitelor *tabele de previziune* în care se înscriu rezultatele evaluării (de pildă notele obţinute la o anumită disciplină sau media de absolvire)) şi rezultatele în integrarea profesională (vezi figura nr. 1).

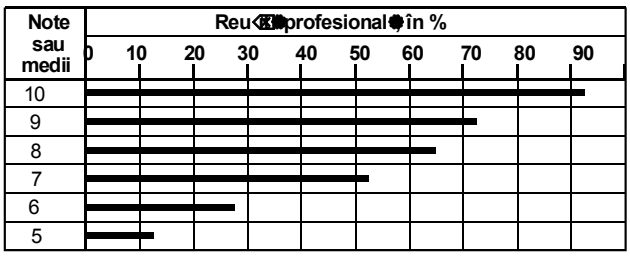


Figura nr. 1. Tabel de previziune

2.2. Fidelitatea notării

Fidelitatea este proprietatea unui sistem de evaluare de *a produce rezultate identice sau echivalente la evaluarea repetată a aceleiaşi caracteristici.* Practic, o notare întruneşte condiţia de fidelitate în următoarele condiţii:

- acelaşi profesor acordă aceleaşi note unor studenţi diferiţi dar care au acelaşi nivel de pregătire;
- aceiaşi studenţi sau studenţi diferiţi dar de acelaşi nivel primesc note echivalente de la profesori diferiţi.

Desigur, este vorba de echivalenţe aproximative, dar ele trebuie asigurate totuşi la un nivel minim la care evaluarea să nu introducă distorsiuni. Când notarea este nefidelă ea capătă un caracter oarecum aleatoriu, fiind dependentă mai mult de contextul în care se face evaluarea, de profesorul care o face sau de studentul care este evaluat, decât de nivelul real al performanţei.

Controlul fidelităţii notării şcolare se poate efectua abordând fidelitatea prin prisma *erorilor de măsură*. Această abordare porneşte de la premisa că, teoretic, nota acordată unui student la verificări diferite trebuie să rămână aceeaşi dacă performanţele studentului nu se schimbă. Această notă, teoretic stabilă, constituie "nota adevărată" a studentului. Practic însă, notele acordate unui student, succesiv de către acelaşi profesor sau simultan de către profesori diferiţi, se abat mai mult sau mai puţin, adică se împrăştie în jurul notei adevărate. În această situaţie, fidelitatea poate fi abordată ca o *problemă de calcul al erorilor*, iar creşterea fidelităţii poate fi asigurată pe căi similare cu cele de reducere a erorilor. În acest sens, se cunoaşte că erorile sunt în funcţie de *N*, adică de *numărul măsurătorilor.* Cu cât *N* este mai mare, cu atât erorile calculate sunt mai mici. Din punct de vedere practic, creşterea numărului de măsurători poate fi făcută de profesor pe două căi:

- *prin creşterea numărului de verificări* ale aceluiaşi student, mai ales atunci când primele evaluări au condus la rezultatele discordante. De exemplu, un student care are o notă de 4 şi una de 9, va trebui reevaluat până când rezultatele vor sugera o tendinţă centrală mai evidentă. Fireşte, este vorba de evaluările continue, din cursul semestrului. Dar această posibilitate este oferită studentului şi în cazul evaluărilor finale prin recunoaşterea dreptului la reexaminare pentru mărirea notei.

- *prin creşterea numărului de întrebări sau a lungimii probei*, având în vedere că orice verificare, probă, test sau chestionar, nu reprezintă decât un fragment limitat al criteriului pe care-l măsoară. Fireşte, procedeele de evaluare nu pot fi extinse peste anumite limite, din motive practice lesne de înţeles: devin greu utilizabile, produc oboseală care influenţează atât performanţa, cât şi motivaţia elevilor.

În tehnica măsurării psihologice se recomandă trei modalităţi de evaluare a fidelităţii instrumentelor de măsură:

a) ***Tehnica aplicării repetate*** constă în aplicarea de două ori a aceluiaşi instrument (de aceea, tehnica se numeşte şi *test-retest*), stabilindu-se măsura în care cele două aplicări conduc la rezultate convergente.

b) ***Tehnica înjumătăţirii*** constă în aplicarea o singură dată a unei probe şi împărţirea ulterioară a probei în două părţi egale, fiecare parte fiind evaluată separat. În final, corespondenţa dintre cele două evaluări va oferi un indice de fidelitate a probei. Este important ca cele două părţi să fie echivalente sau cât mai bine echilibrate ca număr de probleme, ca nivel de dificultate şi ca diversitate de sarcini.

c) ***Tehnica formelor echivalente*** constă în utilizarea a două forme (instrumente) de evaluare pentru evaluarea aceleiaşi caracteristici sau a aceloraşi studenţi (de exemplu, verificarea scrisă combinată cu examinarea orală). Fidelitatea

rezultă din gradul de coerenţă a rezultatelor produse de cele două forme echivalente.

2.3. Sensibilitatea notării

Privită ca o calitate a măsurării, *sensibilitatea este proprietatea unui sistem sau instrument de măsură de a reflecta cât mai fidel variaţia reală a mărimilor caracteristicii măsurate.* Sistemul de notare şcolară poate fi considerat, în anumite limite, un sistem de măsurare, aşa încât el poate şi trebuie să fie analizat din punctul de vedere al condiţiei de sensibilitate. Problema poate fi pusă în felul următor: reuşesc notele acordate să sesizeze corect diferenţele dintre performanţele notate? Răspunsul la această întrebare poate fi dat studiind parametrii ce caracterizează sensibilitatea unei tehnici de măsurare. În teoria şi practica măsurării se consideră că există doi parametri principali ai sensibilităţii:

1. **Pragurile absolute** sunt cele mai mici, respectiv cele mai mari valori ale caracteristicii pe care instrumentul de măsură le sesizează. Diferenţa dintre *pragul maxim* şi cel *minim* reprezintă ***intervalul de variaţie.*** Pentru sistemul de notare şcolară, intervalul de variaţie se situează, teoretic, între 1 şi 10. Practic însă, notele de 1 şi 2 se acordă foarte rar pe criterii de performanţă, aşa încât scara de notare reală se întinde de la 3 la 10.

Problema sensibilităţii, sub aspectul pragurilor absolute, este aceea a definirii, în planul performanţelor studenţilor, a *nivelului* ***maxim*** şi a *nivelului* ***minim***. De modul cum se rezolvă această problemă depind toate celelalte mărimi ce vor fi stabilite în interiorul intervalului de variaţie. Întrucât scara de notare este o *scară ordinală*, oricare dintre nivelurile ei nu are sens decât prin raportare la celelalte niveluri şi, în cele din urmă, la nivelurile absolute, minime şi maxime. Problema este pe ce bază se stabilesc standardele de etalonare a sistemului de notare. Sub acest aspect se pot lua în considerare două tipuri de standarde:

a) ***Standardele generale*** sunt aplicabile întregului registru de performanţe ale studenţilor, la toate disciplinele, cu folosirea unor criterii cum sunt: completitudinea, corectitudinea, frecvenţa şi natura erorilor etc. Standardele generale au menirea de a da semnificaţii comune notelor şcolare, indiferent de disciplina de învăţământ şi de a face notele comparabile. Ele se mai folosesc pentru determinarea nivelurilor minime de promovare, precum şi ca indicatori de eficienţă a procesului de învăţământ.

b) ***Standardele specifice*** sunt proprii fiecărei discipline de învăţământ, deduse în principal din programa şcolară a disciplinei, din obiectivele predării şi învăţării acesteia. Standardele specifice sunt o adaptare a standardelor generale la particularităţile disciplinelor de învăţământ. Fireşte, prin concretizare apar şi criterii de evaluare noi, care nu rezultă nemijlocit din standardele generale. Ele ţin de specificul performanţelor şi al conţinutului fiecărei discipline.

Transpunerea standardelor minime/maxime în etaloane de măsurare şi respectiv în praguri de sensibilitate a notării ridică însă problema raportului dintre nivelurile standardelor prestabilite şi nivelurile reale ale performanţelor studenţilor. De exemplu, într-o grupă sau serie de studenţi cu performanţe excelente, nivelul maximal ar putea fi atins de majoritatea studenţilor (80-90%), ceea ce ar conduce la o notare

nesensibilă (care nu mai sesizează diferenţele dintre studenţi). Această problemă îşi poate fi rezolvată prin utilizarea unor ***standarde normative***. Particularitatea acestora constă în aceea că nivelurile minime şi maxime ale performanţelor corespunzătoare notelor minime şi maxime nu se mai iau ca prestabilite, ci *se stabilesc în funcţie de distribuţia reală a performanţelor în grupul sau formaţia de studenţii în care se face evaluarea*. Procedura de lucru constă în a identifica, de exemplu, într-un set de lucrări, lucrarea cea mai bună şi, respectiv, cea mai slabă, de a stabili ce note corespund celor două lucrări, urmând ca diferenţierea să se realizeze în interiorul intervalului de variaţie astfel stabilit. Desigur, procedând astfel se diminuează comparabilitatea notelor între grupuri de niveluri şi cu distribuţii diferite. Dar această procedură are şi avantajul că verifică şi validează în cele din urmă realismul standardelor generale.

2. **Pragurile relative** sunt cele mai mici diferenţe dintre valorile succesive ale caracteristicii pe care un instrument sau un sistem de măsurare le poate deosebi. În alţi termeni, pragurile relative, numite şi ***diferenţiale***, reprezintă cele mai mici cantităţi care adăugate sau scăzute dintr-o mărime dată conduc la o măsură diferită. Această cerinţă generală este pe deplin aplicabilă şi sistemului de notare. Practic, un sistem de notare nesensibil acordă aceeaşi notă pentru performanţe apropiate, dar diferite. Un sistem sensibil conduce la o notare care deosebeşte între niveluri de performanţă apropiate şi nu include în aceeaşi unitate de măsură mărimi diferite. Această diferenţă minimă pe care sistemul de notare o poate surprinde se numeşte ***intervalul său de rezoluţie***.

În cazul în care notarea este nesensibilă, distribuţia de frecvenţă a notelor acordate este în formă de "I" sau "J", ceea ce înseamnă că elevii clasei nu se distribuie pe o curbă normală, ci sunt plasaţi toţi sau majoritatea, fie în zona notelor mari, fie în zona notelor mici. Această situaţie este datorată în principal diferenţelor mari (în plus sau minus) dintre nivelul de dificultate al sarcinilor şi nivelul real al grupului de studenţi. Când sarcinile sunt prea facile, majoritatea studenţilor se situează într-o zonă restrânsă a scalei de notare (notele 8-10), situaţie redată în figura nr.3. O situaţie analogî apare şi pentru sarcinile prea dificile, cu deosebirea că restrângerea intervalului apare în partea notelor slabe a scalei (notele 4-6), situaţie redată în figura nr.2.

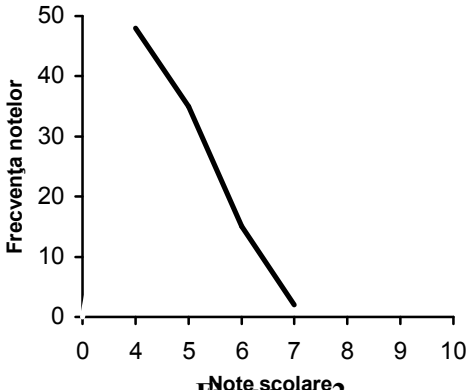


Figura nr. 2

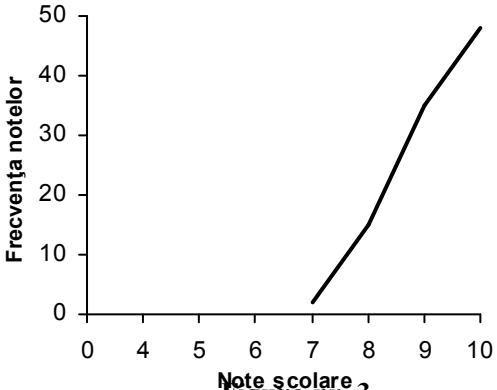


Figura nr. 3

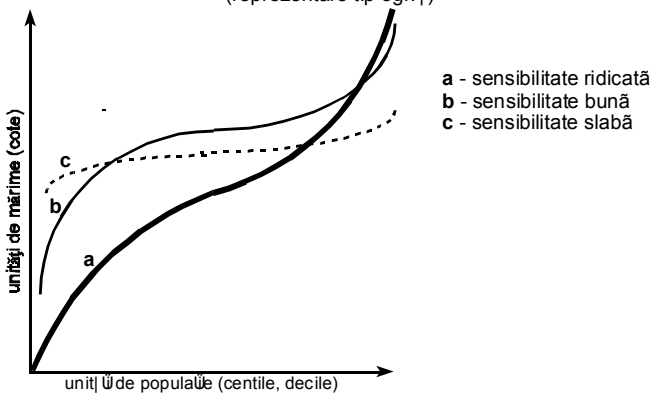
În ambele cazuri, problema care se pune este ***dificultatea medie***, adică acel nivel al dificultăţii în jurul căruia se distribuie, "în jos sau în sus", notele şcolare. Problema dificultăţii medii se pune atât pentru fiecare întrebare, sarcină, problemă, în general pentru fiecare item al unei probe de evaluare, cât şi pe ansamblul solicitărilor pe baza cărora se acordă o notă. Practic, profesorul trebuie să răspundă întrebării: "*Câţi studenţi trebuie să rezolve o problemă sau un item pentru ca acesta să fie reţinut în probă?*". La această problemă s-au dat în timp mai multe soluţii, iar în prezent soluţia cel mai larg acceptată este *distribuţia normală*, potrivit căreia dificultatea medie corespunde itemilor care sunt rezolvaţi de aproximativ 50% dintre studenţii grupului pe care se aplică proba.

Este evident că un instrument (o metodă, o tehnică) de evaluare, pentru a fi sensibilă, nu trebuie să cuprindă doar itemi de dificultate medie, ci şi itemi cu dificultăţi mai mari, respectiv mai mici, într-o proporţie corespunzătoare distribuţiei normale. Ca regulă practică s-a stabilit că procedura cea mai bună este ***creşterea gradulă a dificultăţii***.

Criteriul sensibilităţii criteriului de evaluare nu trebuie însă absolutizat. El trebuie însoţit şi de *criteriul* ***utilităţii***, sau al sensibilităţii necesare. Nu este întotdeauna nevoie să diferenţiem "milimetric" studenţii, ci este suficient să-i ordonăm în grupe de nivel corespunzătoare celor 7-8 niveluri practice ale scării de notare.

Gradul de sensibilitate a notelor acordate de profesor poate fi uşor pus în evidenţă prin stabilirea frecvenţelor de apariţie a notelor şcolare şi prin reprezentarea grafică a distribuţiei acestor frecvenţe. Cu rol orientativ prezentăm în figura nr.4 forma distribuţiei de frecvenţă a notelor caracteristică pentru trei nivele de sensibilitate: ridicată, medie şi slabă. Primul grafic (A) este construit în sistemul decilelor, al doilea grafic (B) în sistemul cotelor normalizate.

A. Pentru sistemul decilelor Ő centilelor
(reprezentare tip ogiv|)



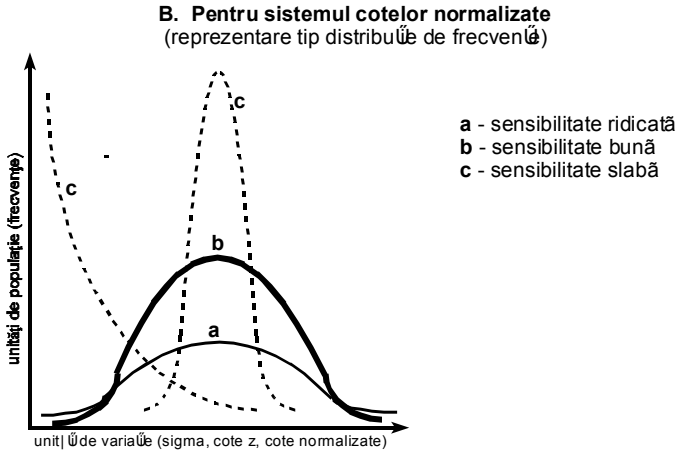


Figura nr. 4

2.4. Obiectivitatea notării

În general, se admite că o *evaluare obiectivă este o evaluare independentă de influenţele subiecţilor implicaţi în realizarea ei (profesori şi studenţi)*. Premisa implicită a acestei afirmaţii este că evaluarea, prin componentele ei de ordin atitudinal şi motivaţional, prin interesele pe care le angajează, ca şi prin complexitatea însăşi a actului de apreciere, poate fi influenţată, în sens pozitiv sau negativ, de subiectivitatea celor care evaluează sau/şi a celor care sunt evaluaţi.

Problema obiectivităţii în evaluare şi, în particular, problema influenţelor subiective în acordarea notelor pot fi sintetizate sub forma a două întrebări:

- *În ce limite intervenţia factorului subiectiv, fie că este vorba de profesor, de elev sau de altă persoană influentă, poate să se abată, în plus sau în minus, de la ceea ce ar fi definit, să spunem, drept o evaluare corectă?*
- *În ce proporţii influenţa factorului subiectiv poate fi distribuită pe cei doi "agenţi" principali: profesorul şi elevul?*

La aceste întrebări, se pot formula următoarele răspunsuri:

1. În primul rând, trebuie subliniat că, oricât de puternică ar fi, intenţionat sau neintenţionat, ***intervenţia factorului subiectiv nu poate depăşi anumite limite ale abaterii de la nivelul corect al evaluării***, chiar şi atunci când acest nivel este cel al "bunului simţ". O performanţă al cărei nivel real este de nota şapte, nu poate fi cotată nici cu zece, dar nici cu patru, chiar şi atunci când evaluatorul ar urmări în mod intenţionat să supraevalueze sau să subevalueze un anumit student. Excepţiile, dacă apar, nu fac decât să confirme regula, căci subiectivismul devine în acest caz evident, iar discuţiile asupra lui trebuie orientate în altă direcţie. Dacă însă renunţăm la prezumţia de intenţionalitate, atunci este cert că *abaterile (erorile) de origine subiectivă nu pot depăşi - şi nici nu depăşesc în plan real - anumite limite.* Tocmai aceste limite sunt cele care trebuie luate în considerare în diminuarea, prin metode şi tehnici adecvate de evaluare, a aportului subiectiv la denaturarea evaluării.

2. În al doilea rând, trebuie să avem în vedere că influenţele subiective pot să aparţină nu numai profesorului, dar şi studentului. Cu alte cuvinte, trebuie să avem în vedere şi măsura şi modul în care ***studentul contribuie, într-un fel sau altul, la influenţarea subiectivă a evaluării***. Concepţia studentului despre evaluare, criteriile proprii, motivaţiile şi atitudinile lui sunt cel puţin la fel de influente ca ale profesorului. În evaluarea prestaţiilor proprii sau pe ale colegilor, studenţii pot comite aceleaşi erori de evaluare pe care le întâlnim şi la profesori. De aceea, formarea *capacităţii de autoevaluare a studenţilor* trebuie să constituie un obiectiv major al procesului de învăţământ, prin care se poate creşte gradul de obiectivitate a evaluării.

3. În al treilea rând, pentru a determina mai bine contribuţia factorului subiectiv în evaluare, trebuie să distingem între două tipuri de influenţe subiective: *o influenţă pozitivă*, care apare ca o consecinţă a angajării subiective a profesorului şi studenţilor în actul evaluării şi *o influenţă negativă*, deformatoare, care produce abateri cu sens de eroare de la caracterul obiectiv al evaluării. Operând cu această distincţie, vom denumi influenţa pozitivă a factorului subiectiv (profesori, studenţi) cu termenul de *subiectivitate*, iar influenţa negativă cu termenul de *subiectivism*. Considerăm că această distincţie creează un cadru mai bun pentru o analiză pertinentă a contribuţiei factorului subiectiv şi, în consecinţă, poate oferi soluţii adecvate pentru controlul influenţei subiective în actul evaluării.

Subiectivismul, definit ca *manifestare negativă a influenţelor subiective*, prezintă fără îndoială un interes mult mai mare decât latura pozitivă a acţiunii factorului subiectiv în actul evaluării. De altfel, majoritatea studiilor referitoare la influenţa factorului subiectiv în evaluare au în vedere tocmai acest gen de influenţă, în acelaşi sens fiind direcţionate şi cele mai multe tehnici de control al intervenţiei factorului subiectiv. În acest sens, considerăm că abordarea cea mai potrivită este de a face distincţie între două tipuri de influenţe negative: *intenţionate* şi *neintenţionate*, aceste tipuri de influenţe fiind semnificative pentru tipurile de efecte generate de distorsiunile subiective: ***erori*** sau ***incorectitudini***.

a) ***Influenţele subiective neintenţionate*** sunt cel mai frecvent întâlnite, iar studiile de evaluare se referă aproape exclusiv la acestea. Este firesc să fie aşa întrucât majoritatea profesorilor nu îşi propun în mod intenţionat să evalueze necorect sau să vicieze rezultatele evaluării. În general, erorile subiective de evaluare îşi au originea în complexitatea acestei activităţi, în gradul mai mic sau mai mare de incertitudine care însoţeşte orice apreciere a calităţilor umane. În particular, distorsiunile subiective de evaluare, pot fi grupate pe cauze specifice, dintre care cele mai importante sunt:

- insuficienţa informaţiilor primare pe baza cărora se efectuează evaluarea;
- metode şi tehnici de verificare şi evaluare inadecvate în raport cu obiectul evaluării (ceea ce se evaluează) şi cu obiectivele evaluării (în ce scop se evaluează);
- influenţele indirecte ale contextului psihosocial în care se efectuează evaluarea (statutul familiei elevului, unele presiuni exercitate indirect de colectivul didactic ş.a.);
- influenţele contextului pedagogic în care se face evaluarea (nivelul general al grupului de studenţi şi compoziţia acestuia, unele elemente de politică a universităţii în materie de evaluare, de pildă, atingerea unei anumite

promovabilităţi, stimularea succesului sau diminuarea insuccesului şcolar al studenţilor etc.).

Efectele influenţelor subiective neintenţionate în evaluare sunt *imperfecţiunile* sau *erorile* de evaluare, ele au *semnificaţie cognitivă şi nu morală*, făcând parte, în anumite limite, din normalitatea actului de evaluare. Existenţa acestor erori este şi motivul pentru care ameliorarea evaluării trebuie să fie o permanenţă a procesului de învăţământ. În acest sens, fiecare din cauzele enunţate solicită căi şi modalităţi specifice de îmbunătăţire a actului de evaluare, aflate în cea mai mare parte la îndemâna profesorului. Sensul diminuării sau înlăturării erorilor de acest gen este cel al *controlului* şi *autocontrolului* influenţelor subiective, prin mai buna cunoaştere şi folosire a metodelor şi tehnicilor de evaluare.

b) În ceea ce priveşte ***influenţele subiective intenţionate***, aceastea au o cauzalitate complexă, inclusiv cu implicaţii social-morale. Problema principală a erorilor generate de influenţele subiective intenţionate constă în aceea că acest tip de influenţe este greu de sesizat şi, mai ales, este greu de demonstrat. Supra sau sub evaluarea intenţionată, cu caracter de avantajare/dezavantajare a unor studenţi, nu apar, de regulă, foarte evident, ele fiind mascate prin folosirea unor criterii, metode şi tehnici "obiective", uneori excesiv de riguroase şi pe fondul unei "imparţialităţi" greu de pus la îndoială. Chiar dacă sunt mai puţin frecvente şi nu pot fi considerate caracteristice pentru profesori, aceste intervenţii subiective există totuşi şi pot avea consecinţe foarte importante, mai ales când evaluarea însăşi are importanţă specială, de pildă când decide obţinerea sau pierderea unei burse sau a statutului de student subvenţionat. De aceea, *corectitudinea morală* şi *respectarea codului deontologic al profesiunii didactice* sunt cerinţe de maximă importanţă ale evaluării şi, în general, ale comportamentului pedagogic al profesorului.

Partea a II-a

EROARILE DE EVALUARE ŞI NOTARE. CUNOAŞTERE ŞI CONTROL

Numim erori de evaluare acea categorie de erori generate de situaţia că între metodele, instrumentele şi tehnicile de evaluare, pe de o parte, şi rezultatele evaluării, pe de altă parte, *se intercalează filtrul subiectiv al persoanei care face evaluarea.* În cazul evaluării şcolare, erorile subiective de evaluare se manifestă practic prin ***distorsionarea relaţiei dintre nivelul real al performanţelor studenţilor şi nivelul calificativului, notei sau punctajelor acordate de profesor, de evaluator***. În fapt, se poate ajunge la situaţia ca nivelul şi dinamica notelor acordate să nu mai reflecte fidel nivelul şi dinamica performanţelor elevilor. Fie că performanţele studenţilor cresc sau descresc, dar evaluările rămân constante, fie că evaluările se modifică fără legătură cu evoluţia reală a performanţelor studenţilor.

Deoarece evaluarea şcolară presupune, explicit sau implicit, efectuarea de *comparaţii* între performanţele unui student şi ale celorlalţi, între performanţele aceluiaşi student la evaluări diferite, precum şi între performanţele studenţilor şi standardele (baremele) de evaluare, erorile subiective de apreciere acţionează cel mai frecvent în *zona estimărilor*, ele viciază *comparaţiile* prin aceea că o estimare deja

făcută influenţează celelalte estimări, prin contagiune, prin contrast, prin inerţie, prin proximitate ş.a.m.d.

Sub alt aspect, erorile subiective de apreciere îşi au originea în implicarea subiectivă a evaluatorului şi evaluatului, implicare generată de *caracterul relaţional,* însoţit de *interacţiuni psihosociale*, al procesului de evaluare. Evaluarea şcolară nu se produce, de regulă, între persoane necunoscute şi nici nu apare ca un act izolat de celelalte interacţiuni ce au loc între profesor şi studenţi în procesul de învăţământ. Pe scurt, relaţia de evaluare este *o relaţie inter-subiectivă* care îşi pune amprenta, într-un fel sau altul, pe rezultatele evaluării. În acest context, vom încerca în cele ce urmează să prezentăm principalele cauze şi principalele efecte ale erorilor de evaluare. Pentru fiecare tip de erori vom încerca să prezentăm modalităţile de identificare şi posibilităţile de control şi de diminuare.

1. Efectul halo şi stereotipiile

Efectul halo şi stereotipiile sunt două moduri de manifestare curentă a erorilor subiective de apreciere. **Efectul halo** constă în tendinţa evaluatorilor de a evalua o anumită caracteristică sau o anumită persoană în funcţie de evaluările făcute anterior, fie că acestea aparţin aceluiaşi evaluator fie că aparţin unor evaluatori diferiţi. Efectul halo conduce practic la o evaluare globală, nediferenţiată, cu o slabă dispersie a calificativelor obţinute de acelaşi individ. În practica evaluării şcolare, efectul halo apare, cel mai frecvent, sub forma *impresiei generale* şi a *globalizării* evaluării.

1.1. Efectul de *impresie generală*

Efectul halo se manifestă cel mai frecvent prin *impresia generală* asupra studentului, formată în funcţie de evaluările anterioare primite de acelaşi student la aceeaşi disciplină sau/şi la celelalte discipline de învăţământ. Această impresie se constituie într-un *clişeu* de apreciere - al profesorului sau împrumutat de la alţi profesori - care poate fi favorabil sau nefavorabil pentru student şi care *maschează diferenţele dintre prestaţiile studentului în cadrul unor evaluări diferite*. Aşa se face că un student cunoscut ca "slab" este cotat cu calificative inferioare chiar şi atunci când prestaţia lui este substanţial mai bună (comparativ cu studenţii buni sau foarte buni), iar un student cunoscut ca "foarte bun" este cotat cu calificative superioare chiar dacă o anumită prestaţie a sa este slabă.

Pe scurt, ***efectul halo se manifestă prin tendinţa evaluărilor de a rămâne constante, chiar dacă nivelul performanţelor evaluate înregistrează variaţii***. Acest efect este cu atât mai puternic, cu cât diferenţa dintre nivelul unei anumite prestaţii a studentului (cea care se evaluează) şi nivelul impresiei generale este mai mare. Astfel, un student care, prin prestaţiile sale anterioare, a creat o impresie generală de nota cinci, va primi cu greu o notă mai mare de şapte, chiar şi atunci când răspunsurile sale sunt bune sau foarte bune (de nota 9 sau 10). Tot astfel, un student cu o impresie generală de nota zece, va primi cu greu o notă mai mică de şapte, chiar dacă un anumit răspuns se situează la nivel de cinci sau şase.

Un mod de manifestare a efectului halo, întâlnit frecvent la profesorii mai puţin siguri pe ei sau la profesorii tineri este de "a arunca o privire" prin catalog sau prin carnetul de note al studentului, înainte de a acorda nota, pentru a vedea "cum stă în

general" studentul la celelalte discipline de învăţământ.

1.2. Globalizarea evaluării

A doua manifestare frecvent întâlnită a efectului halo este *globalizarea evaluării* pe ansamblul criteriilor de evaluare, tendinţa profesorului de a acorda acelaşi calificativ sau calificative foarte apropiate *pentru toate caracteristicile evaluate*, chiar dacă prestaţia studentului nu este egală în raport cu fiecare din criteriile de evaluare. Astfel, dacă un student este considerat superior din punctul de vedere al anumitor caracteristici, de pildă este ordonat, disciplinat, sistematic, riguros, apare tendinţa de a fi evaluat în mod similar şi pentru alte criterii, de pildă pentru originalitate, creativitate sau logica abordării.

Pe scurt, efectul halo apare în acest caz ca ***un transfer nejustificat al evaluărilor pozitive sau negative privind anumite caracteristici în evaluarea pozitivă sau negativă a altor caracteristici***. Sunt implicate aici nu numai distorsiuni de ordin cognitiv sau imperfecţiuni ale tehnicilor de evaluare, dar şi *mentalităţi* şi *atitudini* ale profesorilor - uneori şi ale studentului - privind valoarea, semnificaţia şi importanţa caracteristicilor evaluate. Astfel, pentru unii profesori, comportamentul disciplinat, ordonat, receptiv sau chiar docil al studentului este cotat ca foarte favorabil şi extrapolat în evaluarea celorlalte aspecte ale performanţelor studentului. Se poate întâmpla, de asemenea, ca activismul studentului la seminarii şi cursuri să fie foarte bine apreciat şi extins în evaluarea altor caracteristici.

1.3. Stereotipiile

Stereotipiile sunt forme de manifestare a efectului halo care constau în instalarea unei ***fixităţi a opiniei*** formate despre un anumit student sau grupă de studenţi. Unii autori tratează stereotipia ca pe un factor distinct de efectul halo, considerând că acesta din urmă se deosebeşte prin caracterul său pronunţat afectiv. Se arată în acest sens că stereotipia rezultă dintr-o contaminare a rezultatelor: o primă lucrare mediocră duce la presupunerea că şi a doua va fi la fel de mediocră; dacă aceasta se confirmă, tendinţa de a acorda o notă mediocră şi celei de-a treia lucrări creşte mai mult şi aşa mai departe. Se instalează astfel un *conservatorism cognitiv* care face ca profesorul să nu mai sesizeze schimbările în performanţele şi atitudinile studenţilor.

Ca ilustrare a abaterilor la care poate conduce stereotipia, G. De Landsheere citează cazul unui profesor care a menţinut constantă nota acordată unui elev "cunoscut ca elev slab" chiar şi atunci când, sub numele acelui elev, au fost plasate lucrări ale celui mai bun elev din clasă, apoi ale celui mai bun elev din şcoală şi apoi ale unui licenţiat în domeniul respectiv, dar "nota nu a variat nici măcar cu o jumătate de punct din douăzeci".[99, pag.33].

Trebuie subliniat că efectul halo şi stereotipia pot să acţioneze nu numai asupra evaluărilor diferite privind acelaşi student, dar şi asupra evaluărilor efectuate în *grupe sau serii diferite de studenţi,* în funcţie de climatul general sau de impresia generală pe care fiecare grup de studenţi le-a generat în timp. Astfel, studenţii dintr-o formaţie cotată ca bună, disciplinată, receptivă, harnică etc. tind să fie notaţi mai bine decât studenţii dintr-o altă formaţie care are o imagine nefavorabilă.

Identificarea acţiunii efectului halo şi stereotipiilor se poate face, cel mai simplu,

pe cale statistică, prin analiza datelor şi reprezentarea grafică a distribuţiei notelor acordate aceluiaşi student. Distribuţia tipică pentru acţiunea efectului halo şi a stereotipiei se prezintă ca în figura nr. 5.

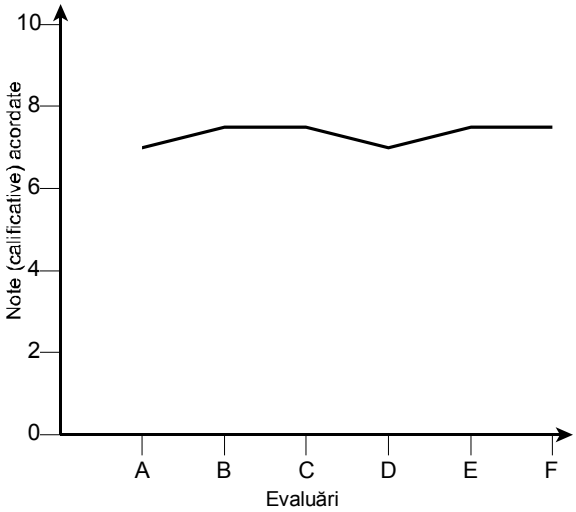


Figura nr. 5 Distribuţia statistică tipică pentru efectul *halo*

Literele (A, B, C, D, E, F etc.) notate pe abscisa graficului pot să desemneze:
- note acordate de acelaşi profesor, la aceeaşi disciplină, aceluiaşi student, în cadrul unor evaluări diferite (prin metode diferite şi în etape diferite);
- note acordate de profesori diferiţi, la aceeaşi disciplină, aceluiaşi student;
- note acordate de profesori diferiţi, la discipline diferite, aceluiaşi student;
- note acordate de acelaşi profesor, la aceeaşi disciplină, aceluiaşi student, dar pentru criterii diferite, de exemplu pentru corectitudine, completitudine, logică, originalitate, calitatea prezentării (scrise sau orale) etc.;

Desigur, graficul trebuie interpretat cu atenţie, mai ales în funcţie de semnificaţia datelor înscrise pe abscisă. Astfel, dacă pe abscisă sunt înscrise notele acordate aceluiaşi student la discipline diferite, *graficul este cu atât mai semnificativ cu cât disciplinele respective sunt "mai diferite"*

Dacă pe abscisă sunt înscrise notele acordate aceluiaşi student, dar pentru criterii diferite, *graficul este cu atât mai semnificativ cu cât caracteristicile evaluate - care constituie criterii de evaluare - sunt "mai diferite" sau divergente.* Astfel, dacă notele sunt aceleaşi pentru două caracteristici cum sunt "tratarea completă a subiectului" şi "abordarea originală a temei", este foarte probabil să avem de-a face cu o evaluare halo, în virtutea ipotezei că o abordare originală nu poate fi, de regulă, şi completă, sau completitudinea nu poate fi pusă în evidenţă deoarece o astfel de abordare "nu se potriveşte" sau "scapă" grilei de evaluare a completitudinii.

În toate situaţiile, măsura în care efectul halo şi stereotipia afectează evaluarea este pusă în evidenţă de forma graficului şi de *gradul de paralelism al liniei frânte cu*

abscisa. Cu cât linia graficului este mai puţin frântă sau cu cât se apropie mai mult de forma unei linii drepte, cu atât efectul halo este mai prezent. Sub celălalt aspect, cu cât linia graficului este "mai paralelă" cu abscisa, cu atât stereotipia este mai pronunţată.

În legătură cu gradul de paralelism al liniei graficului cu abscisa, trebuie spus că, în practică, se pot întâlni încă două situaţii: una în care linia să fie *crescătoare*, cealaltă în care linia să fie *descrescătoare* (vezi figura nr. 6).

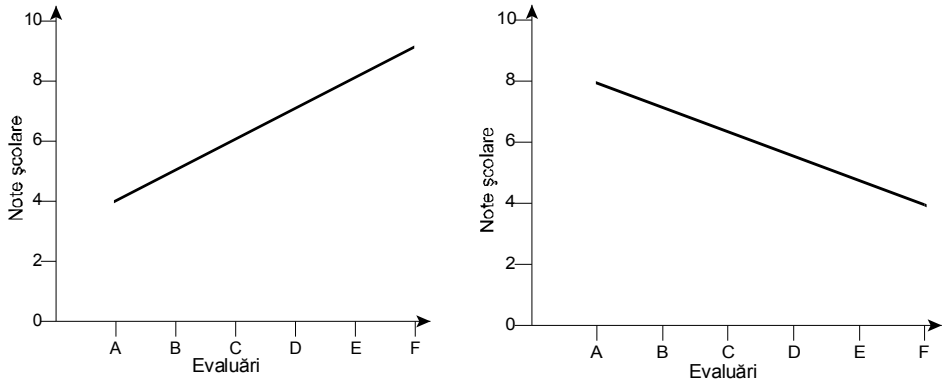


Figura nr. 6. Reprezentarea grafică a efectului *halo* crescător şi descrescător

În prima situaţie, efectul halo acţionează în sensul că unui student care a înregistrat progrese evidente pe parcursul a câtorva evaluări succesive *continuă să i se acorde note mari, în creştere, chiar dacă performanţele lui stagnează sau nu cresc* în acelaşi ritm cu notele acordate. A doua situaţie este inversă: unui student aflat în "cădere", *continuă să i se acorde note tot mai mici, chiar dacă între timp el şi-a ameliorat performanţele.* Cele două situaţii apar şi în cazul în care scăderea sau creşterea performanţelor la unele discipline de învăţământ antrenează *sub sau supra evaluarea studentului la alte discipline.*

1.4. Modalităţile de control

Modalităţile de control al erorilor generate de efectul halo şi de stereotipii se întemeiază pe analiza cauzelor şi a contextului în care apar aceste fenomene. Nu se pot stabili reţete valabile în orice situaţie. Se poate porni totuşi de la câteva particularităţi ale apariţiei şi acţiunii efectului halo şi stereotipiei, care pot sugera şi principalele modalităţi de control.

1. În primul rând, efectul halo, ca şi stereotipia, sunt fenomene de factură *cognitivă*, ele acţionează la nivelul receptării şi prelucrării informaţiei, printr-un proces spontan de selecţie dintre datele actuale a celor care convin unui model sau unei imagini preexistente despre student. Efectul halo şi stereotipia apar din momentul în care imaginea formată despre student nu mai este o reflectare a evoluţiei studentului, ci *este proiectată asupra acestuia*, substituindu-se evoluţiei sale reale.

De aceea, controlul erorilor trebuie să fie un dublu control: un control intern sau

autocontrol, ceea ce înseamnă că profesorul trebuie să aducă în planul analizei conştiente imaginea creată despre fiecare dintre studenţii săi, şi un *control extern*, prin verificarea metodelor şi tehnicilor de evaluare folosite şi, în general, prin verificarea calităţii informaţiilor pe baza cărora se face evaluarea.

2. În al doilea rând, efectul halo şi stereotipia sunt fenomene care apar în condiţiile unor *evaluări repetate*, în cadrul cărora rezultatele evaluărilor anterioare îşi pun amprenta pe evaluările actuale, iar acestea din urmă prescriu un anumit câmp, o anumită prefigurare pentru evaluările viitoare. Rezultă că erorile ţin de *succesiunea evaluărilor* (în ce ordine, la ce interval se fac) şi de *modalităţile* prin care se efectuează fiecare din evaluările succesive (care sunt metodele şi tehnicile folosite, ce obiective, ce conţinuturi şi ce criterii sunt avute în vedere). Sub acest aspect, efectul halo şi stereotipia afectează nu numai rezultatele evaluării (în sensul unor erori de notare), ci şi *procesul evaluării*, adică obiectivele, conţinutul şi tehnicile de evaluare.

De exemplu, unui student cunoscut ca foarte bun, profesorul va continua să îi adreseze întrebări dificile, interesante, de creativitate sau probleme atipice, iar răspunsurile slabe vor fi puse mai curând pe seama dificultăţii problemelor, decât pe seama studentului. Invers, unui student cunoscut ca slab profesorul va continua să îi adreseze întrebări simple, de "încurajare", probleme obişnuite, care nu numai că nu permit elevului să-şi depăşească condiţia, dar ale căror răspunsuri, chiar foarte bune, nu pot fi cotate la niveluri foarte înalte.

Sintetizând modalităţile de control al erorilor provocate de efectul halo şi stereotipie - luând în considerare şi tehnicile de control recomandate în literatura de specialitate - vom prezenta, în cele ce urmează, câteva din modalităţile şi tehnicile de control ce pot fi utilizate în perfecţionarea evaluării, în sensul eliminării dau diminuării efectelor distorsionante de tipul halo-ului sau stereotipiilor:

1. ***Instruirea personalului didactic*** cu privire la existenţa efectului halo şi a stereotipiei, explicarea conţinutului acestora şi a tehnicilor de identificare a existenţei lor. Instruirea evaluatorilor este considerată de mulţi specialişti drept cea mai bună şi la îndemână cale de îmbunătăţire a evaluării. Se arată, în acest sens, că "experimentele legate de diversele metode de evaluare nu fac altceva decât să demonstreze că cea mai potrivită metodă de optimizare a aprecierilor este să instruim cu multă grijă evaluatorii." [H. D. Pitariu, 132, pag.70]. Deseori este suficient ca profesorii să ştie *cum se manifestă* şi *în ce condiţii apar* efectul halo şi stereotipia pentru ca erorile generate de acestea să fie mult diminuate sau eliminate.

2. ***Diversificarea tehnicilor şi procedurilor de evaluare*** sunt modalităţi menite să diminueze *transferul* şi *interferenţa* evaluărilor între caracteristici şi criterii de evaluare diferite. În acest sens, se recomandă:

a) În evaluarea şi notarea lucrărilor scrise, profesorul să evalueze şi să noteze numai un singur aspect (un singur subiect), numai o singură caracteristică dintre cele care constituie criterii de evaluare, urmând ca numai după ce întregul set de lucrări a fost evaluat şi notat la criteriul (subiectul) respectiv să treacă la următorul criteriu (subiect) de evaluare ş.a.m.d. La evaluarea fiecărui nou criteriu (subiect), profesorul nu se va "uita" la notele acordate studenţilor la criteriile (subiectele) anterioare.

b) Combinarea mai multor forme şi tehnici de evaluare (orale, scrise, practice) pentru evaluarea aceleiaşi caracteristici sau a aceluiaşi student (sau grup de studenţi).

3. ***Combinarea, în cadrul aceleiaşi metode (tehnici sau instrument) de evaluare a unor scale de ierarhizare sau de notare cu semnificaţii inverse***, este o modalitate de control ce poate fi folosită practic în următoarele moduri:

a) În unele scale evaluarea să înceapă de la polul favorabil sau nivelul maxim către polul nefavorabil sau nivelul minim, iar în alte scale evaluarea să înceapă cu polul nefavorabil (minim) către polul favorabil (maxim). De exemplu, evaluarea poate să înceapă cu definirea standardelor de nota 10 şi în funcţie de acestea se vor defini treptele inferioare şi, invers, evaluarea poate să înceapă cu definirea standardelor minime de promovare (nota 5) şi în funcţie de acestea se vor defini treptele superioare.
b) În unele scale, polului maxim să-i corespundă nota şcolară minimă, iar polului minim să-i corespundă nota şcolară maximă. De exemplu, se pot lua în considerare numărul şi gravitatea erorilor, polului maxim corespunzându-i cea mai mică notă şi polului minim cea mai mare notă.
c) Să se folosească scale cu sisteme de cotare (punctaje) diferite de scala de notare de la 1 la 10, de pildă cu punctaje care să depăşească 20 sau cu calificative exprimate nominal, urmând ca acordarea notelor şcolare să se facă ulterior evaluării, prin sisteme de convertire a punctajelor sau calificativelor în note şcolare.

4. ***Schimbul interevaluatori*** sau ***evaluarea prin reciprocitate***, este o modalitate de control care, deşi destul de puţin folosită, poate reprezenta o soluţie foarte bună de control al evaluării. Desigur, schimbul interevaluatori se poate realiza între profesori de aceeaşi specialitate sau de specialităţi înrudite, situaţia cea mai favorabilă fiind aceea când aceeaşi disciplină sau discipline foarte apropiate sunt predate de doi sau mai mulţi profesori. Practic, se poate proceda în următoarele moduri:

a) profesorul care predă la seria de studenţi A evaluează la aceeaşi temă seria de studenţi B, iar profesorul care predă la seria de studenţi B evaluează seria de studenţi A, în ambele situaţii conţinutul lucrărilor - întrebările, problemele - fiind stabilit de fiecare profesor pentru seria de studenţi la care predă;
b) profesorul seriei de studenţi A întocmeşte proba de evaluare pentru seria de studenţi B, pe care o corectează atât el cât şi profesorul seriei respective şi, reciproc, profesorul seriei de studenţi B întocmeşte proba de evaluare pentru seria de studenţii A, proba fiind corectată de ambii profesori;
c) profesorul seriei de studenţi A asistă la examinarea orală a seriei de studenţi B şi acordă propriile note şi, reciproc, profesorul seriei de studenţi B asistă la examinarea orală a seriei de studenţi A şi acordă propriile note.

Fireşte, rezultatele obţinute în cadrul acestor evaluări încrucişate pot fi supuse ulterior prelucrărilor statistice, devenind astfel posibilă identificarea cu un grad ridicat de certitudine a eventualelor erori de evaluare.

2. Efectul de contrast şi eroarea proximităţii

Atât efectul de contrast, cât şi eroarea proximităţii reprezintă fenomene ce pot să distorsioneze evaluarea datorită *interacţiunii între evaluările efectuate simultan sau succesiv de către acelaşi evaluator*. Spre deosebire de efectul halo şi stereotipie, care acţionează în sensul unei anumite fixităţi a evaluării, indiferentă la schimbările în plan real caracteristicilor evaluate, efectul de contrast şi de proximitate acţionează în sensul că rezultatele evaluării se modifică, fără ca această modificare să traducă o schimbare în plan real a obiectului/subiectului evaluării. Altfel spus, aceeaşi performanţă, aceeaşi lucrare (test, probă) şi, în general, prestaţii de acelaşi nivel sunt cotate în mod diferit în funcţie de poziţia sau locul ocupat în succesiunea sau în configuraţia evaluărilor efectuate pe un grup de studenţi sau pe un lot de lucrări.

2.1. Efectul de contrast

Efectul de contrast se manifestă practic prin mărirea sau micşorarea notei sau calificativului acordat unui student (sau unei lucrări), în funcţie de calitatea mai bună sau mai slabă a prestaţiei (lucrărilor) celorlalţi studenţi. Se întâmplă astfel ca aceeaşi lucrare sau aceeaşi performanţă să primească note diferite şi anume note mai mari dacă succed unor lucrări sau prestaţii mai slabe, sau note mai mici dacă succed sau se produc în contextul unor lucrări sau prestaţii de nivel înalt. Contrastul poate fi *simultan*, în cazul verificărilor de tip oral, când mai mulţi studenţi sunt verificaţi în acelaşi timp, sau *succesiv*, în cazul lucrărilor scrise, când calitatea lucrărilor corectate anterior (şi notele acordate acestora) influenţează impresia asupra lucrării corectate în prezent şi, desigur, nota acordată acesteia.

Constatăm că eroarea de contrast nu ţine de latura atitudinal-motivaţională a evaluării şi nici de imagini globale pre-existente care să vizeze un anumit student sau o anumită caracteristică evaluată. Ea este o *eroare cognitivă* care se manifestă cel mai frecvent la corectarea în serie a lucrărilor scrise din cadrul examenelor, situaţie în care, de pildă, o lucrare de nivel mediu poate fi apreciată ca slabă dacă este evaluată după o lucrare foarte bună, sau ca bună dacă este evaluată după o lucrare foarte slabă.

2.2. Eroarea proximităţii

Eroarea proximităţii este cauzată de construcţia sistemului (instrumentului) de evaluare şi, în principal, de *ordinea* în care problemele, întrebările, itemii sunt plasaţi sau ordonaţi in instrumentul sau tehnica de evaluare folosite. Rezultă că, în timp ce pentru efectul de contrast cauza se află în succesiunea unor evaluări diferite, în cazul erorilor de proximitate cauza trebuie căutată în interacţiunile dintre itemii aceleiaşi probe, adică ai aceleiaşi evaluări.

Eroarea de proximitate se manifestă în sensul că "dacă un item reprezintă un aspect la care cel notat a fost apreciat ca foarte bun, evaluatorul poate fi influenţat în notarea itemului următor şi chiar a celorlalţi. Uneori poate să apară chiar reacţia inversă, dacă la un aspect aprecierea făcută este favorabilă, la următorul va fi nefavorabilă sau moderată ş.a." [H.D. Pitariu, 132, pag. 77].

2.3. Modalităţi de control

Controlul erorilor de contrast şi de proximitate trebuie să vizeze *modul de organizare a evaluărilor* şi *structura (ordinea) itemilor*, a întrebărilor şi problemelor ce alcătuiesc proba sau lucrarea de evaluare. Rezultă că pentru acest tip de erori, controlul trebuie să se efectueze nu numai în etapa finală, de acordare a notelor, ci *începând cu etapa de proiectare a evaluării*, de construire a instrumentelor şi continuând în etapa de verificare propriu-zisă. Din această perspectivă, propunem în cele ce urmează câteva dintre modalităţile de control eficiente şi uşor de aplicat.

1. Ca şi în cazul efectului halo, una dintre cele mai accesibile şi eficiente căi de control a erorilor de contrast şi proximitate este ***instruirea evaluatorilor*** cu privire la natura şi modul de manifestare a acestui tip de erori. Experienţa instruirii examinatorilor la concursurile de admitere în liceu şi în învăţământul superior a arătat că o bună cunoaştere, de către profesori, a riscurilor de eroare face ca ***autocontrolul*** să devină o modalitate foarte eficientă de limitare a erorilor de evaluare.

2. Pentru *erorile de contrast*, o modalitate utilă de control este ***tehnica comparaţiei pe grupuri de lucrări***, care este o variantă simplificată şi adaptată evaluării şcolare a *tehnicii comparaţiei în perechi* folosită în scalele de evaluare a persoanelor. Această tehnică presupune următoarele operaţii:

a) Se efectuează o primă verificare şi notare a întregului lot de lucrări, după care lucrările sunt grupate în funcţie de notele acordate (lucrări cu note în intervalul 9-10, apoi în intervalul 8-8.99 ş.a.m.d.). Numărul grupelor de lucrări se va stabili astfel încât numărul lucrărilor din cea mai numeroasă grupă (grupă care va fi probabil cea a notelor de mijloc) să nu fie prea mare, dacă e posibil să nu depăşească 15-20 de lucrări. Dacă distribuţia notelor se apropie de o distribuţie normală, atunci va fi necesar ca intervalele de grupare la mijlocul scalei de notare să fie mai mici.

În experienţele realizate de noi am folosit următoarele intervale de grupare:

- grupa 1, intervalul sub nota 5.00;
- grupa 2, intervalul [5-6);
- grupa 3, intervalul [6-7);
- grupa 4, intervalul [7-7.66);
- grupa 5, intervalul [7.66-8.33);
- grupa 6, intervalul [8.33-9.00);
- grupa 7, intervalul [9-10].

b) Se efectuează apoi comparaţii între lucrările din fiecare grupă, prin *tehnica perechilor* (fiecare lucrare este comparată cu fiecare din celelalte lucrări ale grupei) şi, dacă este cazul, se modifică nota acordată iniţial şi se realizează (re)ierarhizarea lucrărilor funcţie de note.

c) Se efectuează *comparaţii între grupe* în modul următor: lucrarea cea mai bună din grupă se compară cu lucrarea cea mai slabă din grupa superioară, iar lucrarea cea mai slabă se compară cu lucrarea cea mai bună din grupa inferioară. În funcţie de rezultatul comparaţiilor, unele lucrări pot să fie transferate în altă grupă (inferioară sau superioară) şi, bineînţeles, re-notate.

3. Controlul *erorilor de proximitate* se poate realiza ***respectând regulile de construire a instrumentelor de evaluare*** (probe, teste, chestionare). Acestea sunt însă

destul de complicate şi nu pot fi aplicate - uneori nici nu este necesar - pe cazul evaluărilor curente din procesul de învăţământ. Se pot utiliza totuşi procedee simplificate, uşor de aplicat şi la îndemâna profesorului.

În general, aceste procedee pornesc de la faptul că erorile de proximitate afectează *fidelitatea* instrumentului de evaluare. Nefidelitatea probelor de evaluare are în principal două cauze: *numărul redus de itemi (de întrebări, de subiecte)* şi *ordinea inadecvată* în care aceştia sunt plasaţi în probă. Cu cât numărul itemilor creşte, fidelitatea probei este mai ridicată, cu precizarea că dincolo de anumite limite creşterea numărului de itemi nu mai aduce o îmbunătăţire substanţială de fidelitate.

Pentru examenele scrise, o tehnică uşor de aplicat este ***procedeul înjumătăţirii***, care constă în divizarea în două părţi egale a probei, fiecare parte fiind corectată şi evaluată separat. De regulă, înjumătăţirea se face luând în prima jumătate itemii impari (1,3,5,7,...), iar în cealaltă jumătate itemii pari (2,4,6,8,...). Nota finală se calculează ca medie aritmetică a celor două note acordate distinct.

3. Ecuaţia personală şi efectul de similaritate

Fenomenele de apreciere subiectivă analizate până acum ţin, în cea mai mare parte, de natura complexă a evaluării, manifestându-se în modalităţi relativ asemănătoare la majoritatea profesorilor, cu diferenţe de nivel de la un profesor la altul, dar fără amprente personale foarte puternice. Spre deosebire de acestea, ecuaţia personală şi efectul de similaritate *sunt specifice fiecărui profesor*, ele sunt mult mai legate de personalitatea profesorului, de concepţia şi atitudinea acestuia faţă de actul evaluării, faţă de procesul de învăţământ în general.

3.1. Ecuaţia personală

Ecuaţia personală este un factor care face ca evaluarea să se diferenţieze de la un profesor la altul, în funcţie de nivelul de exigenţă specific examinatorului şi de concepţia sa cu privire la rolul şi funcţiile notelor şcolare. Practic, intervenţia acestui factor conduce la situaţia că acelaşi nivel de performanţă sau acelaşi student întrunesc evaluări diferite la diferiţi profesori, iar efectul cel mai important este că *evaluările devin necomparabile* şi pot fi avantajoase sau dezavantajoase pentru studenţi. Acest efect devine foarte evident atunci când notele şi mediile obţinute de studenţi au importanţă socială directă (la obţinerea unei burse sau la un concurs de ocupare a unui post). În astfel de situaţii, studentul care a avut "ghinionul" unor profesori exigenţi va fi dezavantajat faţă de ceilalţi.

Ecuaţia personală este un factor care, prin aceea că se menţine relativ constant la acelaşi profesor, nu generează erori propriu-zise de apreciere, ci doar *diferenţe proporţionale ale scării de apreciere* de la un profesor la altul. Sub acest aspect, identificarea ecuaţiilor personale şi atenuarea diferenţelor sunt relativ simplu de realizat, prin *analiza statistică* a evaluărilor şi aplicarea unor *procedee de moderare*. Pe de altă parte, cercetările au relevat că ecuaţia personală a unui profesor nu este întotdeauna constantă, ea poate înregistra variaţii în timp sau de la o categorie de studenţi la alta. Există profesori care sunt mai exigenţi la începutul anului universitar şi mai "generoşi" în evaluările finale. Sub celălalt aspect, unii profesori sunt mai exigenţi faţă de studenţii foarte buni (adică la nivelul notelor mari), în timp ce alţi

profesori sunt mai exigenţi în acordarea notelor mici, de "trecere".

Ecuaţia personală ţine însă nu numai de nivelul de exigenţă specific al profesorului, dar şi de *concepţia sa generală privind nota şcolară*: ca mijloc de *stimulare* sau ca instrument de *sancţiune*. Influenţele acestei concepţii sunt mai evidente în acordarea notelor foarte mari şi a notelor foarte mici. De regulă, profesorii care văd în nota şcolară un mijloc de stimulare, de încurajare, acordă cu mai multă uşurinţă notele mari, ei notează mai mult *progresul* studenţilor decât performanţa absolută, iar nivelul cel mai mic al notelor se află rareori sub patru. Profesorii care consideră nota ca instrument de sancţiune, de avertisment sau ameninţare, acordă în mod frecvent note foarte mici, de trei, doi sau chiar unu şi mai rar note mari.

3.2. Erorile de similaritate

Erorile de similaritate sunt de aceeaşi natură cu erorile de ecuaţie personală, dar îşi au originea nu numai în nivelul propriu de exigenţă sau în viziunea profesorului asupra rolului notelor şcolare, dar şi în trăsăturile sale de personalitate, în *opţiunile profesorului privind valoarea şi semnificaţia comportamentelor şi capacităţilor studenţilor*. Eroarea de similaritate constă în tendinţa evaluatorului de a-i aprecia pe ceilalţi prin *raportare (prin contrast sau asemănare) la propria persoană*. Astfel, dacă profesorul se consideră în primul rând un om ordonat, sistematic, conştiincios va tinde să evalueze studenţii în funcţie de aceste însuşiri. În consecinţă, un student disciplinat va fi cotat pozitiv şi la alte criterii (inteligenţă, creativitate) sau profesorul va fi mai puţin sever faţă de el, în timp ce severitatea sporeşte faţă de studenţii indisciplinaţi.

Efectul raportării la sine poate să acţioneze şi prin raportarea la *experienţa şcolară proprie* a profesorului (explicaţia pe care profesorul o dă succeselor sau insuccesele sale în experienţa de elev şi de student) sau la *experienţa de părinte* (modul în care explică succesele sau insuccesele şcolare ale propriilor copii). S-a constatat, de pildă, că profesorii ai căror copii au fost sau sunt elevi (studenţi) mai puţin disciplinaţi, nonconformişti sau chiar "recalcitranţi" tind să fie mai îngăduitori cu aceste comportamente şi să le considere mai degrabă ca expresii ale inteligenţei, ale creativităţii sau "exuberanţei" proprii vârstei. În acelaşi sens, se poate remarca destul de frecvent mai multă rigoare, intransigenţă sau chiar intoleranţă faţă de abateri la profesorii care nu au trăit şi experienţa de părinte.

Acest tip de influenţe acţionează nu numai în sfera caracterului disciplinat/ indisciplinat, conformist/nonconformist al conduitei studentului, dar şi în sfera evaluării capacităţilor intelectuale, a nivelului de performanţă. Profesorii ai căror copii au înregistrat sistematic succes şcolar şi performanţe de înalt nivel sunt în general mai exigenţi, mai severi, în timp ce profesorii ai căror copii au avut dificultăţi de învăţare şi performanţe mai modeste sunt mai înţelegători, mai puţin severi şi tind să evalueze mai mult *efortul de învăţare*, decât nivelul performanţei.

Atât ecuaţia personală, cât şi efectul de similaritate au, în planul rezultatelor evaluării, consecinţe comune care constau, în esenţă, în *diminuarea comparabilităţii notelor sau calificativelor acordate de profesori diferiţi*. Necomparabilitatea se manifestă pe două planuri:

a) pe planul *nivelului* general al notelor şi calificativelor, prin note diferite

acordate pentru acelaşi nivel de performanţă sau aceeaşi notă pentru niveluri diferite de performanţă;

b) pe planul *criteriilor de evaluare,* prin modificarea ordinii de prioritate sau a ponderii criteriilor de la un profesor la altul, ceea ce conduce, de asemenea, la evaluări diferite pentru aceeaşi performanţă.

3.3. Modalităţi de control

Ca şi în cazul celorlalte erori subiective de apreciere, controlul şi diminuarea erorilor de ecuaţie personală şi de similaritate se efectuează, în primul rând, printr-o bună ***instruire a evaluatorilor***. Autocontrolul subiectiv exercitat de profesor rămâne o cale eficientă de reducere a erorilor de acest tip.

Ca modalitate specifică de control şi corijare a diferenţelor de apreciere inter-evaluatori s-a consacrat ***moderarea***. Aceasta constă într-un ansamblu de procedee care vizează criteriile, metodele şi modul de organizare ale evaluării, menite să atenueze diferenţele dintre examinatori şi să standardizeze rezultatele unor evaluări diferite. G. De Landsheere defineşte moderarea drept "ansamblul măsurilor ce se iau pentru ca notele examenelor interne să fie comparabile şi deci semnificaţia lor să se unifice la nivelul diferitelor clase de acelaşi tip dintr-o şcoală, dintr-un grup de şcoli, din instituţii similare ale unei regiuni sau ale unei ţări."[99, pag. 173].

În esenţă, moderarea implică operaţii prin care se atenuează diferenţele prea mari între notele acordate de către evaluatori diferiţi, aşa încât "ecuaţiile personale" să nu distorsioneze semnificaţia notelor şi să le menţină la un grad acceptabil de comparabilitate. În acest sens se pot utiliza următoarele procedee:

1. ***Procedeul celui de-al treilea examinator*** se poate utiliza atunci când un lot de lucrări este evaluat de doi evaluatori, în mod independent, iar între notele acordate de aceştia aceloraşi lucrări apar diferenţe mai mari de un punct. Cel de-al treilea examinator efectuează, în acest caz, o reverificare a lucrărilor respective, în prezenţa primilor doi examinatori, prin comparare cu criteriile (baremele de notare) şi stabileşte nota definitivă. În acest proces, el se consultă cu fiecare dintre cei doi examinatori, aducând argumente şi încercând să convingă că în cazul lucrărilor respective au apărut abateri de indulgenţă sau severitate nejustificate.

2. ***Procedeul probelor de cunoştinţe standardizate***, cunoscut şi sub denumirea de "sistemul suedez" (întrucât a fost folosit pe scară largă în Suedia), presupune existenţa unui serviciu naţional de evaluare (pentru învăţământul preuniversitar, un astfel de serviciu este constituit în prezent şi în ţara noastră), care propune teste de cunoştinţe, pe discipline, pentru principalele categorii şi niveluri de performanţă prevăzute în programa disciplinelor. În cazul în care un astfel de serviciu nu există la nivel naţional, el poate fi înlocuit cu un *serviciu de evaluare la nivel instituţional* sau *interinstituţional* (între universităţi partenere).

Este de reţinut că aplicarea testelor nu înlocuieşte verificările şi evaluările făcute de profesor, ci serveşte doar ca *sistem de referinţă* pentru compararea între notele acordate de profesor cu standardele instituţionale (naţionale). În acest sens, profesorul calculează media studenţilor evaluaţi în funcţie de notele acordate de el, apoi media în funcţie de rezultatele aplicării testului şi aduce ajustările necesare.

3. ***Procedeul băncilor de întrebări*** (de itemi) este menit să permită atenuarea

efectelor *ecuaţiei personale a profesorilor* - când media generală mai mică sau mai mare a unei serii de studenţi este cauzată de severitatea/indulgenţa profesorului - şi efectele *nivelului seriei de studenţi* - când media generală mai mică sau mai mare a unei serii de studenţi este cauzată de nivelul şi compoziţia seriei respective. În acest scop, nu se mai folosesc probe gata făcute, ci seturi de întrebări (itemi, probleme), grupate pe nivele de dificultate/accesibilitate. Fiecărei întrebări sau seturi de întrebări i se asociază deci un *indice de dificultate* care exprimă proporţia de elevi ce se aşteaptă că vor rezolva corect întrebarea respectivă sau grupul de întrebări. Fiecare profesor alege întrebările adecvate seriei de studenţi la care predă, ţinând seama de nivelul general al seriei, dar şi de compoziţia acesteia (omogenă, heterogenă). În funcţie de nivelurile de dificultate ale itemilor aleşi, se poate calcula un indice mediu de dificultate a probei, ceea ce va permite ca, ulterior, notele acordate studenţilor să poată fi moderate şi interpretate în funcţie de acest indice mediu.

4. ***Procedeul moderării prin raportare la examenele şi concursurile externe*** este menit să compatibilizeze evaluările interne cu evaluările externe, în principal examenele de absolvire, de finalizare a unui ciclu de studii, cu examenele de acces în ciclurile superioare de studii sau cu concursurile de competenţă pentru ocuparea locurilor de muncă. Procedeul constă în analiza comparativă a rezultatelor obţinute de absolvenţi în timpul studiilor şi la examenele de finalizare cu rezultatele obţinute la concursurile de admitere în ciclurile superioare sau la concursurile de ocupare a posturilor. În funcţie de rezultatele acestei analize se face ajustarea criteriilor şi scării de notare în funcţie de rezultatele examenelor şi concursurilor externe.

Ideea este că o notă sau o medie şcolară, mai mare sau mai mică, nu are o însemnătate în sine, ci numai în măsura în care poate garanta studentului sau absolventului o reuşită ulterioară, şcolară sau profesională, cel puţin egală cu nota sau media respectivă. Fireşte, problema nu se pune atât pentru notele mai mici urmate de reuşite mai mari, cât mai ales pentru notele mari şi foarte mari care nu sunt confirmate prin reuşitele ulterioare ale studentului sau absolventului.

4. Eroarea tendinţei centrale şi a restrângerii de rang

Aceste două erori de evaluare traduc un fenomen frecvent întâlnit în tehnicile de evaluare a persoanelor şi se referă la ***tendinţa evaluatorilor de a evita extremele scalei de evaluare***, aşa încât evaluările se concentrează la mijlocul scalei, în jurul valorilor medii.

4.1. Eroarea tendinţei centrale

Eroarea tendinţei centrale poate fi uşor detectată prin construirea grafică a distribuţiei de frecvenţă, forma graficului fiind "ascuţită" şi având o amplitudine redusă (vezi figura nr. 7). Desigur, tendinţa centrală este expresia unor erori de evaluare numai atunci când se produce pe un eşantion (grup de studenţi) care, în mod real, are o compoziţie "normală" (o distribuţie gaussiană) din punctul de vedere al caracteristicilor evaluate. Pentru eşantioanele omogene (, evitarea tendinţei centrale şi "forţarea" distribuţiei normale devin ele însele surse de eroare.

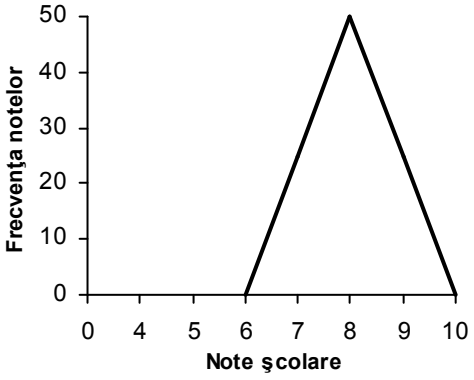


Figura nr. 7 Reprezentarea grafică a unei tendinţe centrale

4.2. Restrângerea de rang

Restrângerea de rang este asemănătoare cu eroarea tendinţei centrale în ceea ce priveşte dispersia evaluărilor, în sensul că amplitudinea variaţiei este mică, dar se deosebeşte prin aceea că media evaluărilor nu se mai situează neapărat la mijlocul scalei de evaluare, ci se poate plasa în orice parte a scalei. Pot să existe restrângeri de rang în zona notelor ridicate, ca şi în zona notelor scăzute. În toate cazurile, restrângerea de rang se manifestă prin numărul mic de poziţii (de ranguri) pe care sunt plasaţi subiecţii evaluaţi, fapt pentru care ea se mai numeşte şi ***eroarea micşorării dispersiei***.

Restrângerea de rang este uneori o formă de manifestare a *efectului halo* şi se datorează imaginii pre-existente despre grupul de studenţi evaluaţi. La un grup de studenţi cunoscut, în general, ca foarte bun, profesorul evită să acorde note slabe, după cum la un grup de studenţi cunoscut, în general, ca foarte slab, profesorul evită să acorde note mari.

Desigur, atât tendinţa centrală, cât şi restrângerea de rang sunt influenţate şi de *ecuaţia personală* a profesorului. De regulă, la profesorii severi apare fenomenul restrângerii de rang în partea de jos a scalei de evaluare, iar la profesorii indulgenţi apare restrângerea de rang în partea superioară a scalei.

5.3. Modalităţi de control

Controlul erorilor tendinţei centrale şi a restrângerii de rang presupune, în primul rând, o *atentă* ***instruire a profesorilor***. Deseori, o bună instruire este suficientă pentru ca acest tip de erori să fie controlat. Foarte important este ca profesorilor să li se sublinieze că tendinţa centrală şi restrângerea de rang apar ca erori numai în raport cu distribuţia reală a studenţilor în raport cu caracteristica evaluată. Dacă grupurile de studenţi sunt omogene, atunci amplitudinea restrânsă a intervalului de variaţie a notelor este o proprietate a grupului de studenţi şi nu a sistemului de evaluare. Sunt utile în acest sens, ***comparaţia interevaluatori***, precum şi folosirea ***evaluărilor prin reciprocitate*** prezentate în paragrafele anterioare.

O tehnică specifică de control al erorilor tendinţei centrale şi restrângerii de rang

este ***tehnica evaluării cu distribuire forţată***. Aceasta constă în a recomanda evaluatorilor să încadreze frecvenţa notelor sau calificativelor acordate în proporţiile distribuţiei normale (gaussiene). În acest scop, se folosesc *scale cu proporţii normalizate* care precizează ce procente din totalul subiecţilor evaluaţi corespund fiecărui nivel al scalei, în condiţiile în care eşantionul evaluat are o distribuţie normală a subiecţilor în funcţie de caracteristica evaluată. Pentru o formaţie cu un efectiv de 100 de studenţi, o scală cu distribuire forţată se prezintă ca în figura nr.8. Pentru efective mai mari (peste 100) se pot folosi scale cu 7 sau 9 trepte, adaptându-se în acest sens proporţiile distribuţiei normale.

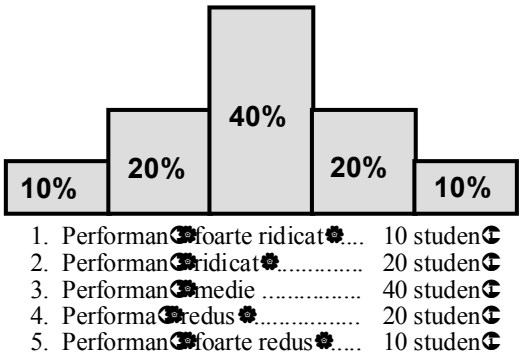


Figura nr. 8 Scală cu distribuire forţată

Deşi valoarea scalei cu distribuţie forţată este deseori contestată, mai ales de adepţii "pedagogiei curbei în J" (o pedagogie a succesului general), folosirea ei poate fi un bun exerciţiu pentru *creşterea sensibilităţii instrumentelor de evaluare* şi aceasta deoarece, cel mai frecvent, eroarea tendinţei centrale şi a restrângerii de rang sunt direct determinate de calitatea metodelor de evaluare. Folosirea unor întrebări şi probleme cu un raport de dificultate/accesibilitate constant nu permite sesizarea diferenţelor reale între nivelele de performanţă a studenţilor şi conduce la o evaluare nediscriminativă. În consecinţă, controlul erorilor tendinţei centrale şi a restrângerii de rang trebuie să se facă nu numai în etapa de evaluare propriu-zisă, de verificare sau corectare, ci şi în etapa de proiectare şi elaborare a instrumentelor e evaluare.

BIBLIOGRAFIE

1. ABRECHT, ROLAND: *L'évaluation formative, une analyse critique,* De Boeck Université, Bruxelles, 1991.
2. BEDARIDA, CATHERINE: *Les examens mal notés*, în Rev. *Le monde de l'éducation*, nr. 204, Mai, 1993.
3. BÂRZEA, CEZAR: *Domenii şi procedee de evaluare*, în vol. *Analiza procesului de învăţământ. Componente şi perspective*, coord. I.Nica, E.D.P., Bucureşti, 1977.
4. CLEMENCE, ALDIN: *Teoriile disonanţei cognitive*, în vol. *Psihologie socială. Aspecte contemporane*, coord. Adrian Neculau, Editura POLIROM, Iaşi, 1996.
5. **** Consiliul Naţional de evaluare şi examinare - Unitatea tranzitorie - *Ghid general de evaluare şi examinare,* Bucureşti, 1995.
6. FIGARI, GERARD: *Évaluer: quel réferentiel?* colecţia *Pedagogies en developpement. Methodologie de la recherche* - De Book Universite, 1994, Wesmael s.a. Bruxelles.

7. HOLBAN, ION: *Testele de cunoştinţe,* E.D.P., Bucureşti, 1995.
8. LANDSHEERE, G. DE: *Evaluarea continuă a elevilor şi examenele. Manual de docimologie* (trad.), E.D.P., Bucureşti, 1975.
9. LISIEVICI, PETRU: *Testele de cunoştinţe. Cadru conceptual, proiectare şi evaluare*, Târgovişte, Institutul de Ştiinţe ale Educaţiei şi CCD, 1993.
10. MĂRGINEAN, ION: *Măsurarea în Sociologie,* Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1982.
11. PAVELCU, VASILE: *Principii de docimologie. Introducere în ştiinţa examinării*, E.D.P., Bucureşti, 1968.
12. PITARIU, HORIA, D.: *Managementul resurselor umane. Măsurarea performanţelor profesionale*, Editura All, Bucureşti, 1994.
13. RADU, IOAN; MICLEA, M; MOLDOVAN, O.; NEMEŞ, S.; SZAMOSKOZY, S.: *Metodologie psihologică şi analiza datelor*, Editura Sincron, Cluj, 1993.
14. VOICULESCU, ELISABETA: *Factorii subiectivi ai evaluării şcolare. Cunoaştere şi control,* Editura Aramis, Bucureeşti, 2001.
15. VOICULESCU, FLOREA; LUDUŞAN, NICOLAE: *Măsurarea şi analiza statistică în ştiinţele educaţiei,* Ed. IMAGO, Sibiu, 1998.