

PROBLEME TEORETICE ALE EVALUĂRII

Dan HUȚĂNAȘU*

Docimologia, pentru a folosi un termen care impune ideea de științificitate, este o ramură tânără a științelor cunoașterii umane. În Franța este recunoscută ca știință după 1920.

Preocuparea pentru evaluare se concentrează, gravitează în jurul unui concept-cheie: *imparțialitatea* sau celebra *obiectivitate*.

Pornind de la această problemă de bază, preocupările docimologilor vor evolua în două direcții:

1) punerea la punct a unor instrumente de măsură (minimizarea variației rezultatelor în funcție de examinatori) și

2) studiul relației dintre formare și evaluare.[5, p. 17] Prima arie de preocupări ia în considerare următoarele aspecte:

- 1) Rolul examinatorului; fiabilitatea testelor
- 2) Lucrările de control sau testele verifică
 - a) achizițiile de cunoștințe sau deprinderi și
 - b) reflectă o anumită filozofie ce ține de conceperea lor, un anumit sistem mai mult sau mai puțin subiectiv ce ține de cel care le concepe.

Nevoia standardizării, validitatea și fiabilitatea sunt concepte de bază care îi preocupă pe cei care concep și aplică testele.

Istoriceste sunt deosebite trei perioade[5, p. 18-20]:

1) "Testing period", pentru a folosi termenul anglo-saxon, care își are rădăcina în testele de inteligență. Binet [5] este cel care concepe primul test în sensul modern al cuvântului. Obiectivul acestuia era unul predictiv: separarea copiilor care puteau urma un program școlar normal de cei cu nevoi speciale.

2) "Measurement period". Este perioada în care se rafinează neconținut principiile testării, în care apar preocupări asupra validității și fidelității, în care se pun bazele prelucrării științifice a rezultatelor.

3) "Evaluation period", care îl are ca figură emblematică pe R. Tyler [5]. Preocupările docimologilor nu mai sunt obligatoriu legate numai de perfecționarea instrumentelor de măsură, ci se ia în considerare întreg procesul de întrepătrundere formare-evaluare. Se iau în

considerare conținutul programelor școlare și adecvarea acestora la grupe de vârstă, validitatea materialului. Testele încetează de a mai fi clasicele lucrări de control, care implică folosirea creionului și a hârtiei. Apar noi instrumente de evaluare: dosarul anecdotic, testele psihologice, testele afective, rapoartele părinților, profilul elevului, aprecieri.

Tot în această perioadă apar și preocupări asupra modului în care intervin și interacționează diferite variabile. Acestea pot interfera semnificativ cu rezultatele școlare, având un impact surprinzător.

Ni s-a părut interesant să prezentăm două dintre multiplele experimente ce pun în valoare modul în care este influențată performanța școlară de identificarea pozitivă/vs./negativă a elevilor.

Conform teoriei lui Mead [3, p. 155-156], imaginea de sine a unei persoane se constituie în cursul procesului de socializare; pentru a putea ajunge la o imagine de sine cristalizată, individul trece prin numeroase interacțiuni, fie ele verbale fie non-verbale [2, p. 155 - 156]. Mead pune accent pe rolul factorului social în eliberarea "conștiinței de sine". În prima fază a constituirii acesteia, sunt interiorizate toate gesturile verbale sau non-verbale ale indivizilor luați separat; în a doua fază, a dezvoltării finale a sinelui, apare ideea socialului, ca individ generic, ca model social, ca grup și nu ca entitate personală.

Prin urmare, socialul intervine hotărâtor asupra procesului de "individuale" și de identificare, folosind orice canal sau mijloc de comunicare. Conceptul de *identificare* reprezintă subscrierea unui individ la un nivel de includere superior, grupul, și este contrariul anonimului. Este în esență reflectată în acest fenomen nevoia acută de relaționare pe care o simțim fiecare dintre noi (*omul comunicant*). Sunt aici celebre experimentele lui Asch [4] asupra conformismului: un individ are permanent tendința să adopte normele, opiniile, atitudinile grupului de apartenență, chiar dacă nu arareori acestea contravin propriilor norme, atitudini, convingeri.[6]

* Asistent universitar, Catedra de Limbi Romanice și Comunicare în Afaceri, Academia de Studii Economice București

Este interesant de constatat în ce măsură *identificarea*, *presiunea socială* și *imaginea de sine* au efecte asupra performanțelor școlare ale unui individ.

S-au făcut experimente numeroase în acest domeniu și toate au dus la aceeași concluzie: identificarea joacă un rol semnificativ în performanțele școlare.

Monteil [3, p.164-166] este unul dintre cei care au condus numeroase cercetări asupra acestei interacțiuni. Unul dintre experimente s-a desfășurat în cadrul cursului de biologie din clasa a șaptea. Experimentul a implicat un grup de școlari cu statut pozitiv și unul cu statut negativ, egale ca număr de membri. Prin *statut pozitiv* autorul înțelege o identificare a elevului cu performanțe școlare superioare, iar prin *statut negativ* o identificare cu rezultate slabe. Această identificare este *exterioară*, adică este operată de grup și de către principalul lui reprezentant, profesorul.

Autorul creează două condiții experimentale: în prima condiție sunt reuniți atât elevi din grupul școlar cu statut superior, cât și elevi din cel cu statut inferior, în număr egal, câte opt. Acestora li se precizează public și nominal statutul. Autorul numește această condiție de *comparație socială*.

În a doua condiție, sunt reuniți elevi din ambele grupuri, tot în număr egal, câte opt din fiecare grup. Li se spune însă că toți au același statut școlar, fără să li se precizeze statutul. Autorul numește această condiție *fără comparație*.

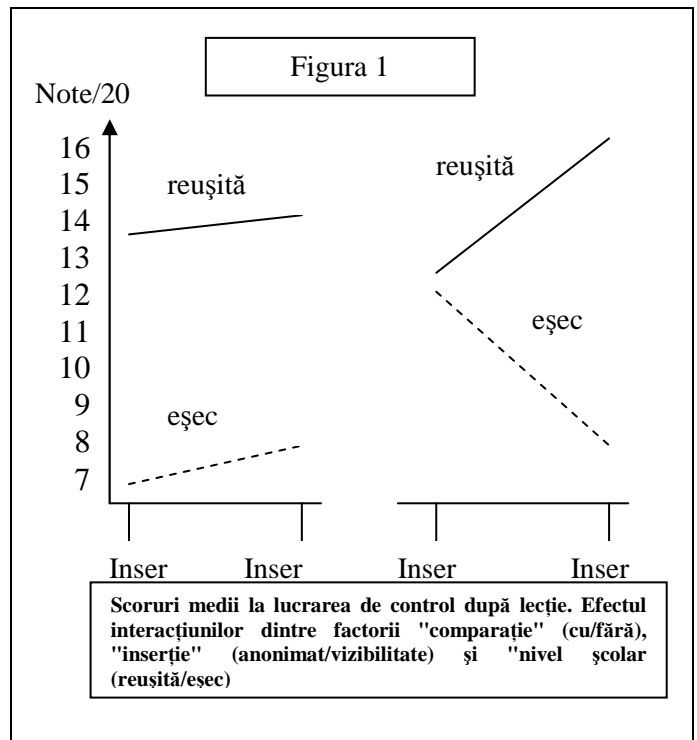
A mai fost introdusă în experiment o variabilă: *inserția*. Înainte de susținerea cursului, unei jumătăți dintre subiecții fiecărei condiții experimentale i se spune că nimănui nu i se vor pune întrebări în timpul expunerii; celeilalte jumătăți i se spune că li se vor pune întrebări pe parcursul expunerii. În realitate, nici-unui grup aparținând celor două condiții experimentale nu i se pun întrebări în timpul cursului. Prima condiție experimentală aparține *unei situații de anonim*, în vreme ce cea de-a doua corespunde *unei situații de vizibilitate*.

Rezultatele experimentului sunt înfățișate în figura 1.

În ambele condiții experimentale, adică în cea *cu comparație* și în cea *fără comparație*, elevii cu statut superior obțin rezultate mai bune decât cei cu statut inferior la testarea de după curs. Cu toate acestea, rezultatele sunt de nuanțat în funcție de a doua variabilă, *inserția*, cu cele două condiții: *anonimat*/vs/*vizibilitate*.

Astfel, în condiția experimentală *cu comparație* rezultatele elevilor cu statut superior sunt net inferioare în condiția *anonimat*, adică atunci când li se spune că nu vor fi întrebări pe parcursul lecției. Dimpotrivă, elevii cu statut superior obțin rezultate semnificativ mai bune în condiția *vizibilitate*, adică atunci când li se

promite că vor fi puse întrebări pe parcursul lecției.



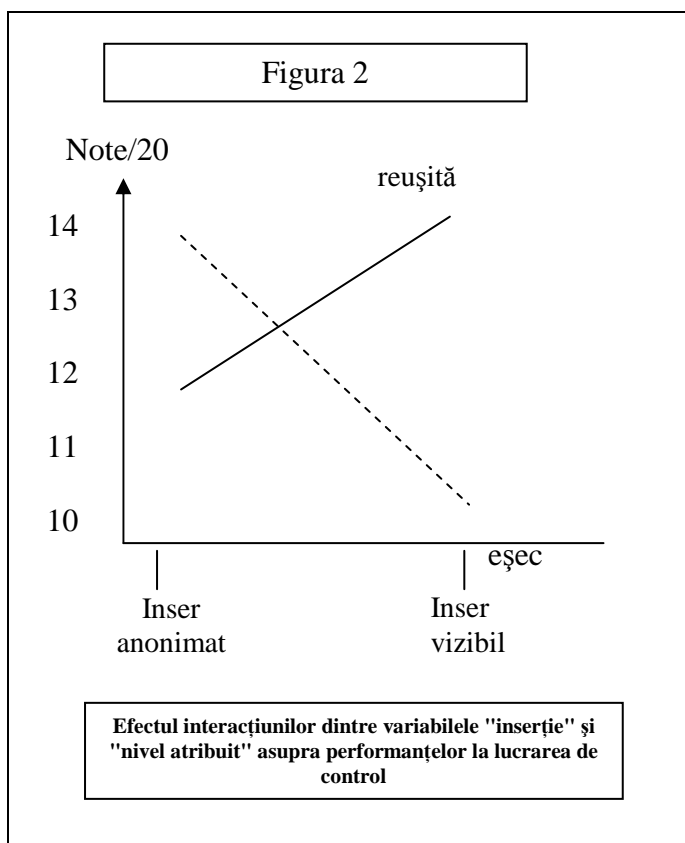
Elevii cu statut inferior obțin rezultate semnificativ mai bune la lucrarea de control de după curs în condiția de *anonimat*, adică fără întrebări pe parcursul lecției. Dimpotrivă, aceștia obțin rezultate semnificativ mai slabe la lucrarea de control în condiția de *vizibilitate*, adică atunci când li se promite întrebări pe parcursul lecției.

Un al doilea set de experimente realizat de Monteil a implicat doar elevii cu statut superior, adică pe cei cu performanțe școlare superioare.

Grupul a fost manipulat altfel decât în experimentul descris mai sus. Jumătate dintre elevi au primit statut inferior, iar cealaltă jumătate statut superior. Această atribuire de statut s-a făcut pe baza unor așa-zise rezultate obținute la o sarcină efectuată cu ajutorul unui microscop. În realitate, atribuirea rezultatelor fiecărui elev în parte era arbitrară, dar se făcea public și nominal. Prin urmare, grupul împărțit în două tabere, cu *statut superior* și cu *statut inferior* se afla în situația de *comparație socială*, variabila independentă fiind constituită din suprapunerea sau nu a statutului avut anterior cu cel atribuit.

A doua variabilă este cea a *inserției*, aceeași ca la experimentul descris mai înainte: unei jumătăți dintre elevi i se spune că i se vor pune întrebări pe parcursul lecției. Aceștia se află în condiția experimentală de *vizibilitate*. Celeilalte jumătăți i se promite că nu vor exista întrebări pe parcursul lecției. Aceștia se află în condiția experimentală de *anonimat*.

Rezultatele celor două grupuri sunt prezentate în figura 2.



Se constată astfel, ca și la experimentul precedent, că situația de *anonimat* este favorabilă la testare celor cu statut inferior atribuit sau nu, în vreme ce situația de *vizibilitate* este favorabilă celor cu statut superior atribuit.

Coroborarea rezultatelor din acest al doilea experiment cu cele din primul întărește convingerea că imaginea de sine și traducerea acesteia în comportamente este determinată social. Astfel, indiferent dacă statutul școlar este sau nu atribuit, elevul tinde să se comporte conform așteptărilor pe care grupul le are față de el. Într-adevăr, deși testarea finală este scrisă, iar lucrările elevilor sunt corectate de patru profesori diferiți, rezultatele diferă semnificativ.

În condiția de vizibilitate socială, performanțele sunt conforme cu sancțiunile atribuite: elevi cărora li se atribuie reușita reușesc, cei cărora li se atribuie nereușita dau greș. Se poate considera astfel că așteptarea examinării orale, care crește pregnanța normelor socio-evaluative ale sistemului școlar, îl determină pe elev să-și înscrie conduita în sistemul de așteptări generate de școală. [5]

Cu alte cuvinte, examinarea orală, care se face în condiții de vizibilitate, adică de față cu ceilalți membri ai grupului, crește predominanța factorilor sociali. Subiectul tinde să-și adapteze conduitele,

performanțele, așteptărilor grupului.

Așa se face că în ambele experimente elevii cu statut școlar inferior, real sau atribuit, au rezultate mai slabe în condiții experimentale de vizibilitate; ei tind să adopte imaginea pe care o are grupul față de ei. Invers, cei cu statut școlar superior, real sau atribuit, au performanțe mai bune, ei tinzând de asemenea să confirme imaginea socială.

Experimentele descrise mai sus se situează în linie dreaptă cu teoria psihologului Merton (1953) asupra profeției auto-îndeplinite. Acesta definește profeția auto-îndeplinită ca fiind: "*la început, o percepție falsă ce provoacă la persoana în cauză un comportament nou care până la urmă ajunge să valideze falsa percepție inițială.*" [3, p.166]

În cele două experimente se poate constata în ce măsură performanțele școlare ale elevilor tind să confirme așteptările grupului. Cel mai evident exemplu al profeției auto-îndeplinite este mai ales al doilea experiment, în care, deși statutul școlar este absolut arbitrar atribuit, elevul tinde să adopte comportamente care până la urmă confirmă acest statut.

Psihologii au mers și mai departe: o persoană își poate comunica, conștient sau inconștient, așteptările unor persoane, care adoptă comportamente ce răspund acestor așteptări: Katzer, Cook și Crouch, (1982) [5].

Această teorie interesantă este cunoscută în educație mai ales sub numele de "*efect Pygmalion*", după cum l-au denumit Rosenthal și Jacobson (1968) [5, p. 52]. Așteptările profesorului determină anumite comportamente din partea elevilor, care tind să le confirme..

Alți autori, Good și Brophy (1984) [5] califică așteptările profesorului ca pe o previziune despre performanțele prezente sau viitoare ale elevului.

În fine, Marc [5] este cel care definește exhaustiv efectul Pygmalion: este vorba despre o situație, un context - acela de formare - , doi actanți - profesorul și elevul - , un set de așteptări - cele ale profesorului față de elev - și un set de răspunsuri - cel al conduitelor didactice ale elevului. Așteptărilor profesorului, care sunt comunicate mai mult sau mai puțin inconștient, le corespunde, după reechilibrări, un set de conduite adecvat din partea elevilor.

Aceste așteptări se formează, în viziunea lui Marc, conform clasicelor de acum constatări ale psihologiei sociale experimentale care se ocupă în cadrul percepției de formarea primelor impresii.

În concluzie, se poate spune că *Pygmalion* are o caracteristică involuntară, inconștientă, non-neglijabilă, a cărei conturare este una dintre marile probleme ale docimologiei și ale pedagogiei.

Un alt mare set de probleme cu care se confruntă docimologia este cel al raporturilor dintre *evaluarea*

formativă și cea normativă.

Istoricește, distincția dintre cele două concepte este relativ recentă și aparține unui american, Scriven, care definește noțiunea de *summative evaluation* după cum urmează:

...evaluare finală și externă. Aceasta constă într-o apreciere de ansamblu a schimbărilor survenite într-o activitate de formare, care permite echipei decizionale exterioare să se pronunțe asupra acesteia. [5, p.59]

În același timp, autorul definește și conceptul de *formative evaluation* după cum urmează:

...evaluare internă, axată pe proces. Constă într-o apreciere globală sau sectorială a schimbărilor care se produc în timpul unei activități de formare, care permite persoanelor responsabile de conducerea activităților de formare să orienteze, să îmbunătățească și să ducă la bun sfârșit această activitate.[id.]

După cum se remarcă din această ultimă definiție, *evaluarea formativă* este un proces continuu, care se bazează pe un permanent feed-back. Deprinderile, cunoștințele sunt cuantificate în comportamente observabile și măsurabile, atât evaluatorul cât și evaluatul având posibilitatea de corectare a erorilor printr-un permanent feed-back. Evaluatorul adaptează conținutul și metodele de predare a cunoștințelor la particularitățile psiho-intelectuale ale evaluatului. Invers, evaluatul are posibilitatea să facă necesarele corecții pe parcursul activității de asimilare de cunoștințe.

Aceste reșezări reciproce sunt esența procesului de evaluare formativă, de unde și denumirea acestuia. Evaluarea formativă ca filozofie și structură merge în paralel cu educația permanentă. Și aceasta din urmă se bazează pe un feed-back permanent, corpul de cunoștințe vechi fiind constant comparat cu altele noi și stabilindu-se reechilibrări periodice.

Evaluarea sumativă, așa cum arată și denumirea, este globală și vizează nu atât reșezări ale cunoștințelor, deprinderilor, cât mai ales tragerea unei linii finale, consecutive achiziționării unui mare corpus de cunoștințe. Ea nu se bazează pe feed-back, pe modificări comportamentale, ci constată, am putea spune, dă un verdict. Evaluarea sumativă sau normativă face apel la evaluatori externi, tinzând spre un maxim de obiectivitate și pe excluderea factorilor personali (relația evaluat-evaluator) ce ar putea influența performanțele evaluatului.

Distincția dintre cele două tipuri de evaluare este istorică, deci împovărată de consecințe. Posibilitatea corectărilor de parcurs influențează întreaga pedagogie. Fiecare etapă a formării trebuie să conțină obiective precise, cuantificabile, măsurabile și care trebuie atinse până la sfârșitul etapei. Scopul evaluării formative este de a constata atingerea sau nu a acestor obiective de

parcurs. Acestea pot fi însă uneori atât de minuțios circumscrise, încât există pericolul *atomizării* lor. Nu ne putem abține să nu facem, împreună cu Meyer [5, p. 57], o referință istorică: gesturile repetate până la absurd ale muncitorului întruchipat de Charlie Chaplin, care satirizau taylorismul atât de în vogă în America de atunci. Într-adevăr există acel pericol al *taylorizării* învățământului: obiectivele pot fi foarte clare și precise, dar minuscule, așa încât putem pierde din vedere ansamblul, care poate fi ceva compozit, format din gesturi un pic absurde:

Cât despre dascălul de rând, acesta, muncitor specializat la o bandă rulată tayloristă, are grijă ca șuruburile a căror strângere intră în sarcina lui să fie înfiletate corect; el nu are niciodată o viziune de ansamblu asupra produsului pe care-l fabrică sau asupra modului în care se îmbină diferitele părți pentru a forma un întreg, adică un elev competent, după cum nu are nici sentimentul responsabilității în elaborarea acestui întreg. [id.]

Cu toate riscurile ei, distincția *formativ/normativ*, operată la sfârșitul anilor 60 de către americanul Scriven, face carieră, influențând într-un mod sau altul diferitele concepții actuale asupra docimologiei.

Boniol și Nunziati introduc o inovație modificând termenul de *evaluare formativă* în *evaluare formatoare*. [6]

Viziunea autorilor este tehnicistă: există un caiet de sarcini (programele școlare oficiale); un produs (comportamentul măsurabil al elevului); profesorul devine un soi de *birou de studii* ce stabilește un proiect tehnologic; elevul devine la rândul său un soi de *birou de metode*, stabilind el însuși etape; cea mai des folosită metodă este cea a încercării și erorii, de sorginte behavioristă.

Trebuie să constatăm că acești doi autori au imprimat mai ales un limbaj tehnicist în pedagogie și docimologie: lor le *datorăm* faimosul *proiect de tehnologie didactică* în locul clasicului *plan de lecție*. De asemenea, izul său behaviorist limitează serios această concepție. Metoda încercărilor și a erorilor nu poate fi singura în orientarea procesului de învățare și nici în obținerea feed-backului, deși este adevărat că ocupă un rol major.

O altă perspectivă interesantă asupra evaluării este cea a lui Allal [1]. Acesta introduce un triptic: *evaluare prognostică, formativă și sumativă*.

Evaluarea prognostică are rolul de stabilire a direcției, a *itinerarului* pedagogic, în vreme ce celelalte două își păstrează sensurile deja acordate de Scriven. Toate cele trei componente ale evaluării interacționează în cadrul unui model. Astfel și evaluarea formativă și cea sumativă utilizează *instrumente criteriale*. Evaluarea prognostică, împreună cu cea normativă utilizează și *instrumente normative*.

Acest model interacționist vizează rafinarea instrumentelor de măsură, prin multiplicarea interacțiunilor dintre acestea și introducerea lor într-un sistem evaluativ. Prin urmare se încearcă ocolirea criticilor la adresa evaluării, care pretind că evaluarea ar pierde din vedere aportul cunoașterii intuitive. Într-un fel este vorba despre vechea dispută din psihologie între *insight* (intuiție) și *testing* (de sorginte mai mult sau mai puțin behavioristă).

De asemenea, modelul interacționist al lui Allal mai introduce în cadrul evaluării formative o distincție, anume aceea de *reglare integrată* sau *integrativă* vs. *reglare amânată*.

În cazul *reglării integrate*, modificările aduse de elev metodei sale de învățare sunt o consecință imediată a interacțiunilor sale cu profesorul, cu colegii, cu materialul didactic... Reglarea interactivă se sprijină, în cele mai multe cazuri, pe demersuri informale de evaluare reciprocă. Instrumentul ciclului evaluare /reglare îl constituie, în primul rând, structurarea lecției, consemnele oferite elevilor, exigențele impuse temelor, resursele aflate la dispoziție, modul de organizare sau de funcționare avut, acceptabil sau respins, etc. [id., p 35] Este cu alte cuvinte vorba de o panoplie de instrumente ale micii evaluări, mergând de la simplul ați înțeles? care cere un feed-back, până la instrumente diagnostice mai complicate cum ar fi micro-testările, verificarea temelor pentru acasă sau exercițiile cu cheie.

Un teren fertil de investigare în acest domeniu l-ar putea constitui ceea ce am denumi *evaluarea non-verbală*, ce ar ține tot de efectul *Pygmalion*. Profesorul are mai mult sau mai puțin inconștient gesturi sau expresii mimice/pantomimice care reprezintă, în opinia noastră o componentă interesantă a procesului de formare.

Tonalitatea vocii, o anume expresie facială reprezintă un feed-back pe care elevul îl recepționează imediat și pe care îl decodează la modul *bravo, așa este!* sau *cum poți spune o asemenea tâmpenie?* Ponderele acestor indici în procesul comunicării, adecvarea lor la context, simpla lor absență sau prezență pot constitui arii de cercetare viitoare pentru pedagogi.

Al doilea tip de reglare în viziunea lui Allal, cea *amânată*, intervine la sfârșitul unei evaluări sistematice inserate la un moment dat în derularea unei secvențe de activități de formare. Acțiunea de reglare este deci amânată în raport cu lecția/secvența de lecții și în raport cu momentul evaluării... Utilizăm expresia de reglare proactivă pentru a desemna operațiunea de planificare a unor activități ulterioare, orientate mai curând către consolidarea și aprofundarea competențelor elevilor, decât către remedierea dificultăților specifice întâmpinate sau a

greșelilor comise... Prin reglare retroactivă înțelegem punerea în practică a unor activități care îi permit elevului să depășească dificultățile întâmpinate sau să corecteze greșelile reperate prin evaluare... Reglările amânate, retroactive sau proactive, rezultă, în cele mai multe cazuri, din aplicarea unui instrument sau a unei proceduri de evaluare sistematică. [id.]

În fine, sunt de menționat încă doi autori, J: Ardoino și G. Berger [2], care introduc o nouă distincție, aceea dintre *evaluare* și *control*. Această distincție însă, în opinia noastră, seamănă suspect de bine cu cea operată deja de Scriven. *Evaluarea*, în viziunea acestor autori, este structurantă în procesul de vehiculare a cunoștințelor. Folosindu-se de un mic joc de cuvinte, aceștia afirmă:

Informația crucială [în cazul evaluării – n.n.] constă în ceea ce vine în întâmpinarea nevoilor sistemului. [5, p.34]. Cu alte cuvinte, evaluarea are un rol diagnostic, orientându-se către repararea breșelor în sistem. Controlul măsoară abaterea de la conformitatea cu niște norme prestabilite; pentru a folosi o expresie a lui G. Meyer, el este în afara timpului, autoarea punând accent pe obiectivitatea și perenitatea normelor.

În ciuda acestor rafinări de sens, concluzia rămâne, după părerea noastră: evaluarea normativă este egală cu controlul, iar cea formativă, singurul tip de evaluare în viziunea autorilor, nu este alta decât cea descrisă de Scriven.

Înșiși autorii se situează implicit pe această poziție, recunoscând că în spatele evaluării sumative externe, denumire comodă după 1968, se ascunde de fapt un control, care trebuie să-și asume și să-și îndeplinească rolul. [5, p.35].

Cu o privire retrospectivă asupra problematicii ridicate de docimologie, se impun, după părerea noastră, două direcții mari de cercetare în această știință.

Prima mare direcție este cea a studiului *efectului Pygmalion*, deja dovedit. Unde și cât poate repara evaluatorul daunele produse de propriile preconcepții asupra elevilor? În ce măsură presiunea grupului poate fi folosită atunci când este canalizată spre scopuri pozitive și în ce măsură poate fi contracarată atunci când imaginea, identitatea de sine a unui membru îi limitează performanțele? Experimentele lui Monteil ridică problema îmbinării cât mai juste a formelor de testare în funcție de statutul elevilor: *testare scrisă*/vs/*testare orală*.

Al doilea mare grup de probleme este suscitată de clasică distincție dintre normativ și formativ. Ca și la grupul de probleme precedent, se iscă întrebarea în ce măsură putem *tempera*, adică amesteca în proporții juste cele două tipuri de evaluare. De asemenea, nu putem evita un alt semn de întrebare: cum ocolim atomizarea obiectivelor, cu alte cuvinte, în ce măsură

putem intra în amănunte pentru a nu pierde din vedere ansamblul. La câteva decenii deci de la distincția operată de Scriven, rămâne în vigoare dilema legitimității celor două tipuri de evaluare.

Este puțin probabil să se găsească sentințe definitive,

dar aceste probleme ne zgândăresc pentru perfecționarea evaluării. Totodată, în prezent, accentul pare să se deplaseze de pe *formare*, care era *regina* procesului învățării, pe *evaluare*, a cărei însemnătate nu mai poate fi neglijată.

REFERINȚE BIBLIOGRAFICE

1. ALLAL - *Assurer la réussite des apprentissages*, M. Huberman, Delachaux-Niestlé 1988.
2. ARDOINO, J. BERGER, G. - *D'une évaluation en miettes à une évaluation en actes*, Paris, Andsha, 1989.
3. DOISE, W., (coord) - *Psychologie sociale expérimentale*, Armand Colin, Paris, 1991.
4. DROZDA-SENKOWSKA E. - *Psychologie sociale expérimentale*, Paris, Armand Colin, 1999.
5. MEYER, G. - *De ce și cum evaluăm*, Polirom, "Collegium", Iași, 2000.
6. NUNZIATI, G. - *Evaluation formatrice*, in Cahiers pédagogiques, nr. 280, ian. 1990.