

Automatizarea procesului de evaluare a testelor grila

Asist. Cristian USCATU, analist Gabriel SUTAC
Catedra de Informatica Economica, A.S.E. Bucuresti

Evaluarea în învățământ este un subiect controversat datorită, în special, încărcăturii sale morale – prin evaluare se realizează clasificări și selecții, se dau “verdicte” și se hotărăște soarta unor oameni. Adesea este confundată cu verificarea cunostintelor și atribuirea de note pentru acestea. Evaluarea este însă un proces mult mai complex, care presupune un efort continuu de diminuare a subiectivismului din aprecierile profesorilor în acordarea de “note” (calificative, punctaje etc.). Testele grila par să fie metoda de evaluare prin care subiectivismul este eliminat în cel mai mare grad. Problema, în folosirea acestui tip de testare, constă în timpul mare necesar pentru stabilirea rezultatelor și posibilitatea apariției unor greșeli datorate factorului uman. Automatizarea procesului de stabilire a rezultatelor apare ca un obiectiv firesc, cu consecința imediată a creșterii fidelității rezultatelor obținute și minimizării timpului consumat.

Cuvinte cheie: evaluare, subiectivism, grila, scanner, automatizare, software.

Evaluarea în învățământ îndeplinește mai multe funcții, cu efecte în plan individual și social, după cum urmează:

- Funcția de control, de constatare și apreciere a activității și rezultatelor obținute în procesul de învățământ, prin care se stabilește unde se situează aceste rezultate în raport cu obiectivele proiectate, încercând să “depisteze” factorii care influențează în sens pozitiv sau negativ aprecierea. Prin exercitarea acestei funcții, evaluarea are și rol de feed-back.
- Funcția de reglare a sistemului de evaluare, de ameliorare a activității și de optimizare a rezultatelor, care constă în demersurile comune ale evaluatorilor și evaluatilor pentru a face “corecțiile” necesare – pe baza controlului și aprecierilor – în stilul de conducere, respectiv în activitatea de execuție. În cazul evaluării rezultatelor obținute de subiecți, accentul cade pe optimizarea stilului de predare, a capacității de învățare și pe stimularea factorilor motivaționali.
- Funcția de predicție, de pronosticare și orientare, prin care se încearcă prefigurarea desfășurării activității în sistem (sau într-un subsistem) și de anticipare a rezultatelor, ca urmare a măsurilor preconizate.

- Funcția de clasificare și selecție, în baza căreia se ierarhizează instituțiile de învățământ, elevii și studenții.
- Funcția educativă, menită să constientizeze, să motiveze și să stimuleze interesul pentru studiu continuu, pentru perfecționare și pentru obținerea unor performanțe cât mai înalte.
- Funcția socială, prin care se realizează informarea colectivității locale, a familiei, asupra rezultatelor obținute de subiecți. În practică, aceste funcții se întrepătrund, iar în raport cu scopul evaluării, unele vor avea o pondere mai mare decât altele sau se vor urmări în exclusivitate.

Necesitatea folosirii testelor grila

În procesul de evaluare apar factori perturbatori ai aprecierilor, generați de activitatea și trăsăturile de personalitate ale cadrului didactic.

1. *Efectul halo* constă în supraaprecierea sau subaprecierea rezultatelor unui subiect sub influența impresiei generale pe care și-a făcut-o în timp profesorul despre acesta. În acest caz, există tendința de a trece cu vederea unele greșeli sau rezultate mai slabe la subiecții cu o bună reputație și de a ignora progresele școlare sau răspunsurile

excelente date de așa-numitii “elevi-problema”.

2. *Efectul blând* se exprimă în tendința de a aprecia cu indulgență persoanele cunoscute, în comparație cu cele mai puțin cunoscute, profesorul dovedind o mai mare precizie în notare abia după mai bună cunoașterea a tuturor.

3. *Eroarea de generozitate* constă în acordarea unor note mai mari în raport cu nivelul de pregătire a subiecților prin probe deosebit de ușoare și o exigență scăzută. Această atitudine se produce, de regulă, în cazurile în care profesorul este interesat să escamoteze realitatea datorită esecurilor în demersurile sale instructiv-educative, să prezinte unele rezultate superioare pentru a fi apreciat în calitate de cadru didactic sau când se manifestă o empatie deosebită, care, în final, duce la o demotivare a elevilor. De asemenea, efectul de generozitate poate fi produs și de modul simplist de apreciere a activității cadrelor didactice numai prin prisma notelor obținute de elevi la materia respectivă.

4. *Efectul Pygmalion* (sau *efectul oedipian*) are ca efect modificarea comportamentului elevului în raport de convingerea profesorului că acel elev nu poate satisface cerințele școlare. Acest comportament indus elevului conduce, în final, la esec.

5. *Efectul de contaminare* se referă la situația în care cunoașterea notelor atribuite de ceilalți profesori influențează aprecierea unui evaluator.

6. *Efectul de contrast* (sau *efectul de ordine*) constă în mărirea diferențelor în aprecierea unor răspunsuri succesive. Astfel, dacă după un elev care a răspuns foarte bine, urmează un răspuns bun, există tendința de a-l subevalua pe cel din urmă. Reciproca este valabilă, adică un răspuns bun, de regulă, este supraevaluat dacă succede unui răspuns foarte slab.

7. *Ecuatia personală a examinatorului* (sau *eroarea individuală constantă*) rezultă dintr-o exigență specifică fiecărui cadru didactic în aprecierea rezultatelor școlare.

8. *Eroarea de tendință centrală* se concretizează în renunțarea acordării de note foarte mari sau foarte mici elevilor, din precauția de a nu greși sau din dorința de a mulțumi pe toată lumea.

9. *Eroarea logică* presupune înlocuirea unor criterii de apreciere a rezultatelor școlare în raport de obiectivele didactice, ca variabile adiacente acestora: forma grafică, acuratețea lucrării, modalitatea inedită de prezentare etc.

10. *Efectul curbei lui Gauss* rezultă din dorința cadrului didactic ca distribuția rezultatelor școlare în fiecare clasă să fie cât mai aproape de curba lui Gauss. Aceasta presupune ajustarea exigenței în notare în raport de nivelul general atins de fiecare clasă, astfel încât distribuția rezultatelor să reflecte, pe cât posibil, o distribuție normală. Astfel, la o clasă foarte bună exigența crește, iar la alta, cu rezultate în general mai slabe, exigența scade. În acest mod, la cele două clase de niveluri generale diferite, există 2-3 elevi foarte buni și 2-3 elevi foarte slabi.

11. *Factorii de personalitate ai cadrului didactic* pot influența modul de evaluare didactică. Astfel, atitudinea cadrului didactic pe timpul verificărilor, echilibrul și constanța în comportament față de studenți, empatia - ca o componentă a aptitudinii didactice, contribuie la crearea sau nu a condițiilor propice unei evaluări obiective.

Existența acestor factori este cunoscută sau percepută de cei evaluați, ceea ce duce la scăderea încrederii în acuratețea evaluării și diminuarea, în consecință, a calității procesului didactic.

Eliminarea lor în totalitate nu este posibilă în cazul evaluatorului uman, impunându-se în ultimii ani evaluarea prin teste grila. În cazul testelor evaluatorul nu are altceva de făcut decât să constate alegerea făcută de evaluat dintr-o mulțime de posibilități și să acorde punctele corespunzătoare opțiunii.

Erorile care pot apărea tin de factorul uman, dar sunt de alta natură. În funcție de

modul concret de lucru, pot apărea erori la transcrierea rezultatelor, la punerea în corespondența a opțiunilor cu punctajele sau la calculul punctajului total al fiecărui test. O soluție folosită pe larg în trecut a constat în introducerea redundanței: toate operațiunile erau efectuate de două ori, de persoane diferite. Ca efect, a crescut fiabilitatea întregului proces dar și timpul consumat și stresul asupra personalului. Fiind vorba despre activități de rutină, care nu implică inteligența, ele pot fi automatizate, prin folosirea tehnicii de calcul.

Sistem de evaluare automată a grilelor

O variantă de sistem de testare automată este testarea on-line a studenților, în lume existând astfel de sisteme, cu evaluare chiar prin Internet (*Brainbench*). Pentru folosirea acestei variante este necesară o infrastructură tehnică adecvată, care să permită accesul candidaților la sistemul de evaluare în timp real și în paralel. În condițiile actuale din România implementarea unui astfel de sistem presupune cheltuieli care nu pot fi, încă, suportate. Ceea ce se poate face însă este automatizarea etapei de evaluare a grilelor scrise. În acest scop este nevoie de un scanner, care să transpună în format electronic foile cu grilele candidaților, o componentă software care să "identifice" răspunsurile date de aceștia (inclusiv cazurile în care nu s-a dat nici un răspuns ori s-au selectat mai multe răspunsuri), o componentă care să compare răspunsurile cu baremul și să contorizeze punctajul pentru fiecare test, o componentă care să asigure imprimarea rezultatelor evaluării.

Sistemul trebuie să fie tolerant la erorile de identificare: hîrtia pe care sunt imprimate grilele poate să prezinte imperfecțiuni; modul de completare a grilelor de către candidați poate fi defectuos (depășirea zonei desemnate pentru marcarea unui răspuns). Datorită necesității de secretizare a grilelor (lipirea colțului cu datele de identificare a autorului), rămîne o etapă care trebuie

efectuată manual: deschiderea tezelor, punerea în corespondența a tezelor cu listele de punctaje obținute și afișarea rezultatelor.

Avantajele acestui sistem sunt: timpul scurt necesar pentru întregul proces (de la încheierea examenului și colectarea grilelor pînă la afișarea rezultatelor), acuratețea rezultatelor afișate (acestea reflectă exact răspunsurile date de candidați la test), reducerea stresului, atît asupra candidaților care așteaptă rezultatele cît și asupra examinatorilor.

În cadrul Academiei de Studii Economice a fost pus la punct și folosit cu succes un sistem pentru examenele "masive" (admitere și licență, unde numărul candidaților este de ordinul miilor la fiecare facultate). Pentru realizarea sistemului au fost necesare următoarele resurse:

- scanere format A4 cu rezoluție de minim 300dpi. Ideale ar fi scanere cu posibilitatea de alimentare automată a foilor cu teste, care necesită o intervenție minimă din partea operatorului. Se pot folosi și scanere cu alimentare manuală, cu un operator care să asigure alimentarea;
- pachetul Remark Office pentru identificarea răspunsurilor din imaginile scanate ale grilelor;
- un program realizat prin eforturi proprii pentru tratarea cazurilor de erori de scanare/identificare, introducerea punctajelor și calculul notelor. Prin eroare de identificare se înțelege imposibilitatea pachetului de programe de a discerne care varianta de răspuns a fost aleasă de candidat dintre cele posibile (nu sunt erori de identificare a cazurilor în care nu a fost marcat nici un răspuns ori au fost marcate mai multe). Programul utilizează o bază de date creată în acest scop;
- imprimarea rezultatelor se face prin intermediul Microsoft Office.

Metodologia de evaluare a grilelor

Metodologia propusă prezintă cîteva variante, care sunt alese în funcție de configura-

ratia calculatorului folosit, numarul de grile care se scaneaza, calitatea hârtiei pe care se afla grilele si continutul testelor (grila pura sau combinatie: o parte grila si una sau doua probleme care se corecteaza clasic, înainte analizei grilei). Se tine cont de faptul ca fiecare grila este cuprinsa într-o pagina A4 si o întrebare din grila are numai un raspuns corect. Grilele au fost concepute pentru maxim 45 de întrebări pentru fiecare test.

Metodologia aleasa contine urmatoarele etape:

- Se grupeaza foile dupa numarul testului si se numeroteaza secvential fiecare pachet;

- Lansarea programului Remark Office pentru: crearea modelului testului pentru analiza (definirea zonelor care vor fi scanate si analizate); scanarea baremelor de corectare (grile corect completate de catre comisie); scanarea imaginilor, pe rând, sub controlul pachetului Remark (în cazul unei erori de recunoastere se anuleaza pagina);

salvarea rezultatelor într-o tabela a bazei de date;

- Lansarea programului propriu pentru: inserarea manuala a paginilor cu erori de recunoastere care au fost anulate; numerotarea seturilor de raspunsuri; verificarea cazurilor de raspunsuri multiple sau vide, prin confruntare cu testul; introducerea punctajelor pentru fiecare întrebare; introducerea punctajelor la probleme; corectarea testelor prin comparare cu baremul corespunzator si calculul punctajelor; calculul notei finale prin luarea în considerare a punctajelor la întrebări, probleme si puncte din oficiu;

- Imprimarea listei cu rezultate care va contine, pentru fiecare test, numarul de ordine din pachet, numarul de puncte si nota finala;

- Deschiderea tezelor si scrierea listei finale cu rezultate (nume si nota) fac parte din metodologia propriu-zisa, fiind executate manual.

Schema fluxului de date este redată în figura 1:

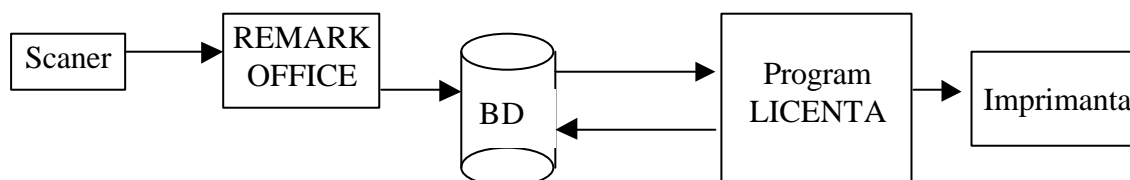


Fig. 1. Fluxul de date al prelucrării automate a grilelor

În timpul folosirii pachetului Remark pot apărea erori de recunoastere a raspunsului. Acestea pot fi tratate prin anularea întregii pagini (urmând a fi introdusa manual în programul propriu) sau prin introducerea manuala a raspunsului corect. În figura 2 se prezinta cazul unei grile introdusa gresit în scanner.

Programul propriu pentru corectarea testelor

Programul *Licenta* este folosit pentru prelucrările specifice, care nu sunt acoperite de pachetul Remark Office. Pentru folosirea acestuia, baza de date trebuie înregistrată ca sursa de date a sistemului în

pagina *Start > Settings > Control Panel > ODBC Data sources > System DSN*.

Elementele de interes ale programului sunt jurnalul activitatilor realizate în baza de date si meniul *Tasks*, care contine optiuni corespunzatoare etapelor descrise anterior (figura 3).

Setarea codurilor tezelor presupune punerea în corespondenta a numerelor de ordine ale tezelor scanate cu seturile de raspunsuri identificate si salvate în baza de date. Numerele se asigneaza secvential, începând cu numarul primei teze din pachet. Vor fi exceptate numerele corespunzatoare paginilor care au fost anulate datorita erorilor de identificare.

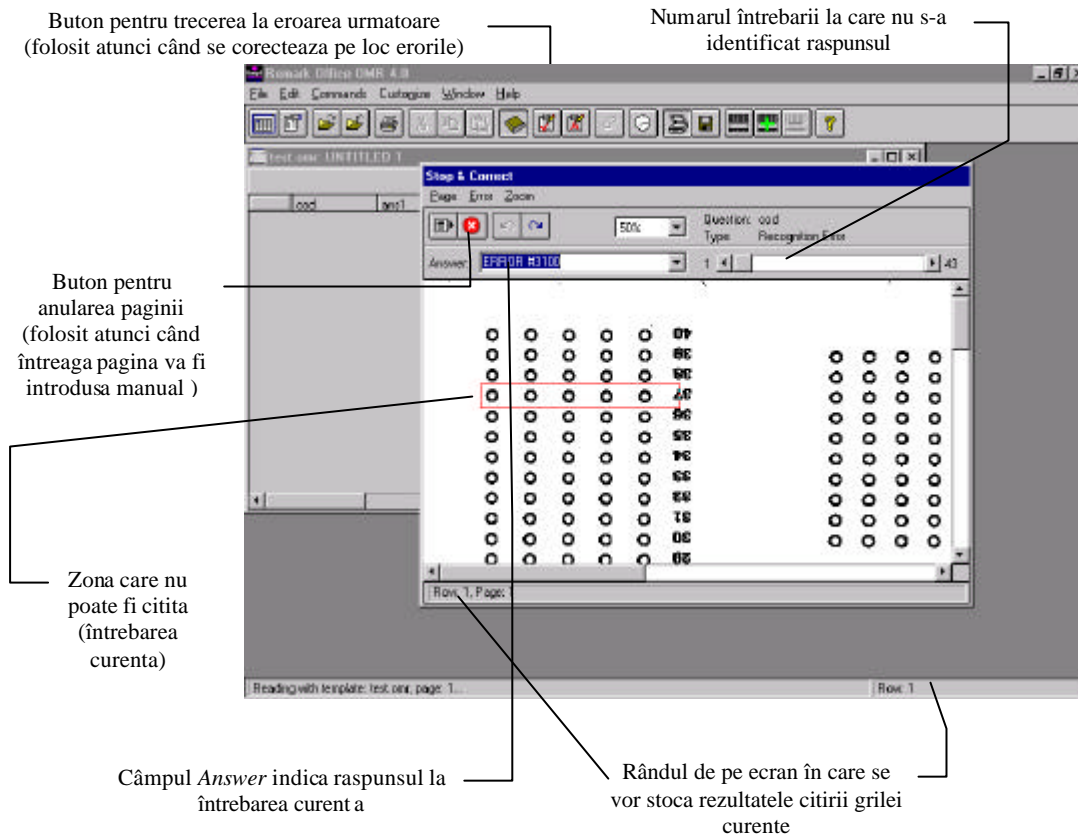


Fig. 2. Semnalarea unei erori de identificare

Meniul *Tasks*

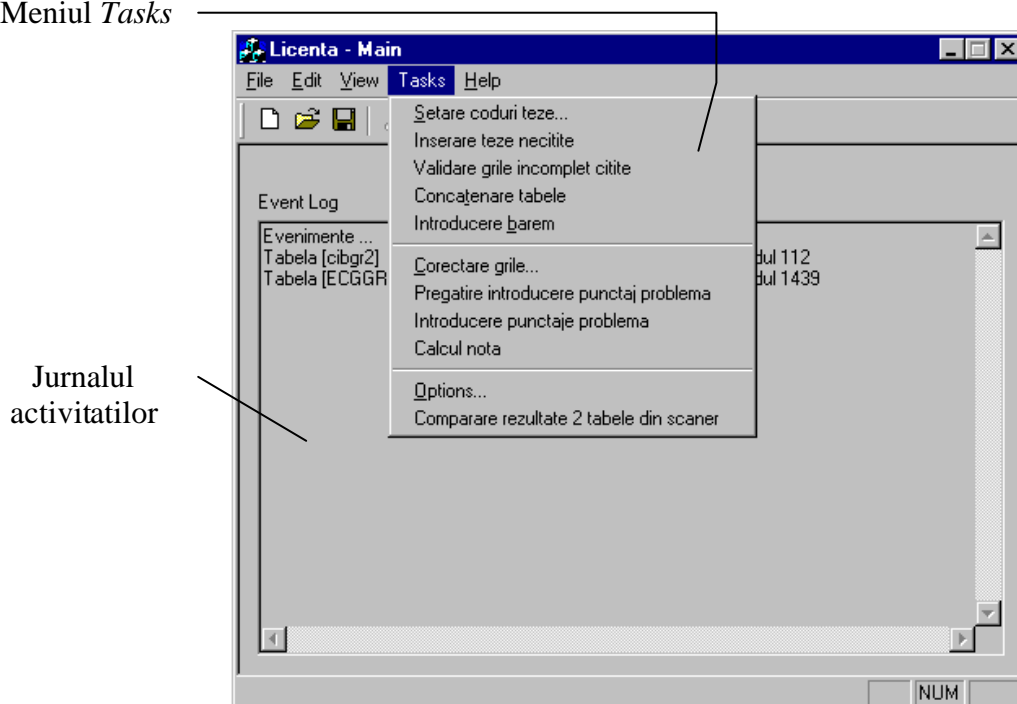


Fig. 3. Fereastra principala a programului pentru calculul notelor

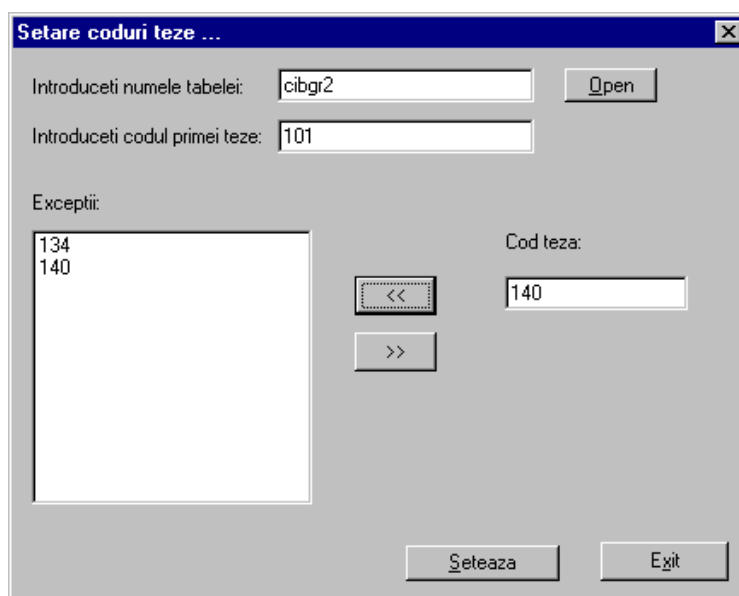


Fig. 4. Setarea codurilor tezelor

Pentru paginile care au fost anulate, se vor introduce manual raspunsurile (*Inserare teze necitite*). Se vor introduce numarul de ordine al tezei în pachetul de teze, codul

disciplinei si codul grilei, formatul grilei fiind comun pentru mai multe examene – discipline (figura 5).

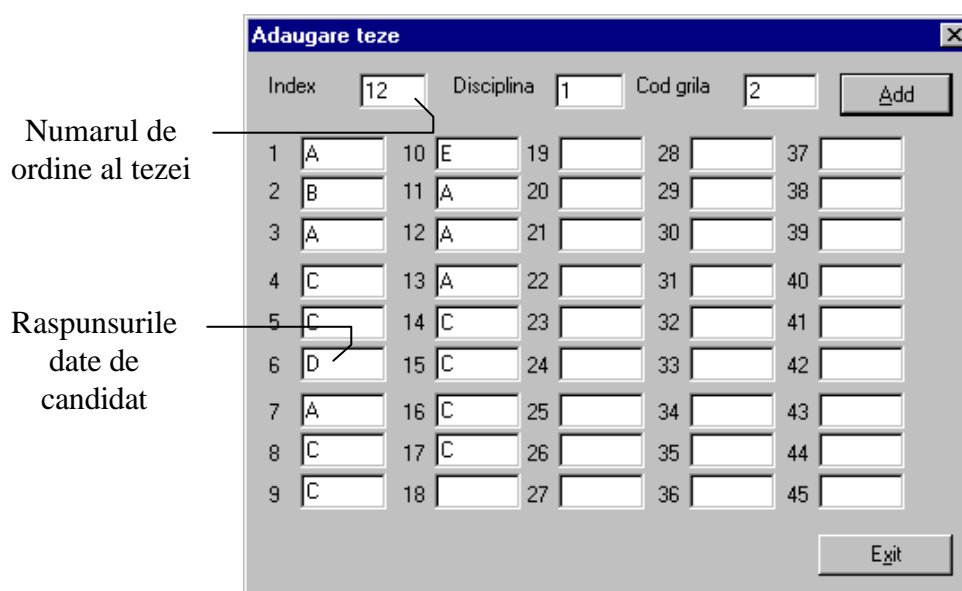


Fig. 5. Introducerea manuala a unei teze care nu a putut fi scanata

În cazul raspunsurilor vide (identificate cu *Blank*) sau multiple (identificate cu *Mult*) se face o verificare suplimentara prin confruntare cu tezele (optiunea validare grile incomplet citite). Programul selecteaza se-

turile de raspunsuri care contin astfel de identificari si le afiseaza, pe rând, în vederea confruntarii si eventualei modificari (figura 6).

Fig. 6. Validarea raspunsurilor multiple sau vide

Pentru calculul punctajului fiecărei teze este necesara introducerea punctajului fie-

cărei întrebări din test prin opțiunea *Introducere barem* (figura 7).

Fig. 7. Introducerea punctajului aferent fiecărei întrebări

Programul calculează în permanentă suma punctajelor introduse pentru verificarea rapidă a corectitudinii datelor introduse. În cazul în care testele conțin și probleme, corectate separat, vor fi introduse punctajele corespunzătoare obținute de candidați

pentru a fi luate în considerare la calculul notelor finale (opțiunile *Pregătire introducere punctaj problema* și *Introducere punctaje problema*). Notele se calculează prin opțiunea *Corectare grile*.

Rezultate obtinute

Metodologia prezentata a fost folosita experimental la sesiunea de licenta din mai 1999, la sectia de Informatica Economica din cadrul Facultatii de Cibernetica, Statistica si Informatica Economica, unde s-au obtinut rezultate foarte bune. Ca urmare, a fost adoptata ca metodologie standard pentru sesiunea de licenta din septembrie 1999, la nivelul A.S.E. si, de asemenea, la concursul de admitere în A.S.E din septembrie 1999. S-a continuat cu sesiunea de licenta din februarie 2000 si, datorita avantajelor, a fost solicitata si în cadrul testelor curente din timpul anului (de exemplu, la *Bazele programarii calculatoarelor*, anul I, Facultatea C.S.I.E.).

În urma folosirii acestei metodologii sa redus drastic timpul consumat între terminarea examenului si afisarea rezultatelor (de la 2-3 zile la câteva ore, rezultatele fiind afisate chiar în ziua examenului). Calitatea rezultatelor a crescut corespunzator înregistrându-se foarte putine erori în calculul notelor finale. Cauzele erorilor tin exclusiv de factorul uman si nu de meto-

dologie. S-au produs erori datorita “murdaririi” tezelor în timpul manipularii lor si datorita introducerii gresite de catre operatorul uman a rezultatelor unor teze, care nu au putut fi scanate (au fost eliminate si aceste erori, prin eliminarea necesitatii de a introduce manual teze). Pe masura acumularii de experienta, s-a ajuns la eliminarea totala a erorilor de acest tip la ultimele examene.

Bibliografie

Remark Office OMR 4.0 - Manualul produsului;

Bull Glen L. – *Technology and Schools*, articol în *ADVANCE IN COMPUTERS*, vol. 45, Academic Press Ltd., 1997;

Coverly C. D. – *Technology in Developmental Education: Past, Present, Future*, Southwest Texas State University, *Proceedings of the Sixteenth Annual Institute for Learning Assistance Professionals*, University Learning Center, University of Arizona, <http://www.schooledu.swt.edu/Dev.ed/Technology/PastPresFuture.html>;

Microsoft Visual C - Manualul produsului