

TESTAREA
STANDARDIZATĂ
ÎN EDUCAȚIE

Dragoș Iliescu, specialist în psihologie aplicată în context ocupațional și în probleme de capital uman, precum și în testarea și evaluarea psihologică și educațională (cu o componentă interculturală puternică), este profesor la Facultatea de Psihologie și Științele Educației din cadrul Universității din București, unde conduce și școala doctorală – și, de asemenea, „profesor extraordinar în cadrul Universității din Stellenbosch. Pe lângă cariera academică, a fost activ în consultanță de-a lungul ultimilor 25 de ani, fiind implicat în sau conducând un număr de proiecte importante privind testarea și evaluarea psihologică și educațională, în contexte comerciale sau guvernamentale, în Europa, Asia, Africa și America de Sud. Dragoș Iliescu a servit în diverse roluri oficiale pentru un număr de asociații profesionale naționale și internaționale; este fost președinte al International Test Commission (ITC) și actualul președinte al Diviziei 2 (Evaluare psihologică și educațională) al Asociației Internaționale de Psihologie Aplicată. Totodată, este redactor-șef al *European Journal of Psychological Assessment* și autor a peste 200 de articole științifice, capitole de carte și cărți, fiind, printre altele, coordonator, împreună cu alți colegi, al volumului premiat internațional *International Handbook of Testing and Assessment*, publicat în 2016 la Oxford University Press, precum și autorul monografiei *Adapting Tests in Linguistic and Cultural Situations*, publicată în 2018 la Cambridge University Press.

DRAGOȘ ILIESCU
TESTAREA
STANDARDIZATĂ
ÎN EDUCAȚIE

Prefață
de Daniel David

 HUMANITAS
BUCUREȘTI

Redactor: Cristian Negoia
Coperta: Ioana Nedelcu
DTP: Dragoș Dumitrescu, Veronica Dinu

Tipărit la Artprint

© HUMANITAS, 2023

Descrierea CIP a Bibliotecii Naționale a României
Iliescu, Dragoș
Testarea standardizată în educație / Dragoș Iliescu;
pref. de Daniel David. – București: Humanitas, 2023
Conține bibliografie
ISBN 978-973-50-8149-2
I. David, Daniel (pref.)

37

EDITURA HUMANITAS
Piața Presei Libere 1, 013701 București, România
tel. 021.408.83.50, fax 021.408.83.51
www.humanitas.ro

Comenzi online: www.libhumanitas.ro
Comenzi prin e-mail: vanzari@libhumanitas.ro
Comenzi telefonice: 0723.684.194

Cuprins

<i>O lucrare în spirit iluminist modern, în contextul citizen science</i> (de Daniel David).....	7
<i>Cuvânt înainte</i>	9
1. Introducere în evaluarea standardizată	13
2. Domenii ale evaluării standardizate	51
3. Efecte pozitive ale evaluării standardizate	96
4. Riscuri, efecte perverse (și alte preocupări legitime) legate de evaluarea standardizată.	128
5. Dezvoltarea testelor standardizate și asigurarea calității.	172
6. Alternative la testarea standardizată.	254
<i>Bibliografie selectivă</i>	285

O lucrare în spirit iluminist modern, în contextul citizen science

Știința a apărut ca demers de sine stătător în perioada modernă, fiind organizată în instituții (academii/universități/asociații/reviste etc.) aflate în relații strânse unele cu altele. Sigur că demersurile științifice pot fi identificate încă din Antichitate/Evul Mediu, dar știința ca instituție care se ocupă de adevăr a fost creată de civilizația noastră modernă, pentru ca oamenii să se protejeze de ei înșiși în căutarea adevărului, dincolo de „tribunalele populare“ sau de regula celui care este mai puternic.

În perioada actuală vedem însă un atac la știință, început cu postmodernismul și accentuat de emanciparea individuală prost înțeleasă în contextul tehnologiilor și rețelelor sociale, în care oamenii confundă înțelegerea proprie cu adevărul/cunoașterea științifică. Răspunsul științei la aceste presiuni au fost demersurile de tip *citizen science*, prin care cunoașterea trebuie dusă la oameni, și oamenii trebuie aduși la știință!

În domeniul abordat de această carte, știința face tot mai greu față pseudoștiinței și simțului comun. Oamenii, în vârtejul dorinței de a-și vedea vocea ascultată, exagerează și nu mai știu când să tacă reflexiv: la modul eronat considerăm că, fiind oameni toți, ne pricepem la oameni și la comportamentul lor, și toți știm cum ar trebui făcută educația, doar am trecut prin asta și/sau avem copii care trec. Când nu iese cum dorim în evaluare/testare, metoda este proastă, nu noi avem ceva de îmbunătățit în ceea ce ne privește. Oamenii nu mai știu când să pună întrebări experților sau cum să formuleze opinii informate. Așa cum spuneam mai sus, gândurile, ideile, înțelegerile noastre sunt prea ușor considerate deja adevăr: sigur că ceea ce este adevărat poate fi înțeles și ne oferă apoi sens și semnificații, dar nu tot ce poate fi înțeles și ne dă sens și semnificații este și adevărat!

Lucrarea de față se înscrie perfect și la timp în această paradigmă a vremurilor noastre, venind atât ca un antidot la problemele cu care se confruntă domeniul și societatea, cât și ca o cale de urmat. În primul rând este o lucrare științifică, scrisă de probabil cel mai titrat psiholog român în domeniu – fost președinte (2016–2018) al International Test Commission (ITC), redactor-șef

al revistei europene de referință (*European Journal of Psychological Assessment*) și autor al unor monografii internaționale devenite un reper în domeniu –, care abordează riguros concepte psihologice și educaționale complexe (testare, evaluare, standardizare, competențe etc.). Dar este o lucrare științifică scrisă într-un limbaj accesibil pentru a ajunge la oameni, în vechea logică iluministă de *Sapere Aude!* Și, pentru a întregi paradigma, lucrarea este scrisă într-un limbaj atractiv, pentru a aduce oamenii cu plăcere și curiozitate la știință.

Daniel David

Cuvânt înainte

Mă bucur că pot oferi cititorului interesat acest volum, care vine să ocupe un gol de informație relevantă despre un subiect care a fost prezent în agenda publică, sau cel puțin în cea referitoare la educație, pentru o perioadă. Și, așa cum se întâmplă de regulă într-o Românie cu discursuri și argumente atât de polarizate, din păcate nu prin dezbateri lucide și așezate, ci prin scrisori deschise, formulate răstit și ultimativ, de oameni care clamează expertiză în domenii la care, de fapt, rareori se pricep mai mult decât tangențial.

Totuși, tema testării standardizate în educație e importantă, și consider că pot propune ceva nou în această discuție. Sper să nu fiu înțeles greșit: îl parafrazez aici pe Nichita Stănescu, care spunea că el crede cu tărie în vocația lui de poet – la fel de puternic cred și eu în vocația mea legată de acest domeniu al testării. Am avut însă oarece dubii în privința priceperii mele de a face inteligibil cititorului interesat, dar nu neapărat și specializat, un domeniu care nu este (trebuie să recunosc asta, comparativ cu domenii ale științelor „grele”) chiar atât de complicat, dar care nu este cu siguranță nici chiar simplu.

Așadar sper să mă achit de acest obiectiv cu decență: să pun la îndemâna celor interesați o carte care prezintă viziunea mea despre testarea standardizată în educație – o carte ușor de citit și de înțeles, care poate fi baza unor discuții publice mai așezate și care, de asemenea, poate fi fundamentul unor procese de învățare sau transformare pentru cei ce o citesc: părinți, profesori, jurnaliști, responsabili de politici publice sau, de ce nu, elevi.

Trebuie să mărturisesc că este prima carte pe care o scriu vreodată fără note de subsol, trimiteri bibliografice și sute de referințe. Acesta a fost planul inițial, este adevărat – și la fel de adevărat este că a fost nevoie de insistența combinată a colegilor mei și a editorului meu pentru a mă motiva să mă țin de acest plan și pentru a șlefui, acolo unde nu m-am putut abține la prima scriere a textului, aceste apucături de scriere științifică.

E posibil să fiu acuzat, pe partea cealaltă, de simplificare excesivă și de tratament neștiințific. Legat de primul punct – simplificarea excesivă –, trag nădejde că nu am căzut în această capcană: da, am încercat să simplific niște subiecte destul de grele, dar sper că nu am simplificat excesiv. Legat de al doilea posibil reproș: nu am avut nici o clipă intenția să scriu o carte științifică legată de acest domeniu – poate va veni și acel moment, căci România duce lipsă de specialiști în testare, evaluare, măsurare educațională, are nevoie de ei, dar nu-i are, sau nu-i are la nivelul de excelență la care se joacă acest joc pe plan internațional. În mod cert însă, nu aceasta a fost intenția aici: acesta e un text simplu, introductiv, chiar dacă poate ușor militant pe alocuri (da, cred în valoarea testării standardizate și în capacitatea ei de a contribui la procesul de vindecare a educației din România).

Mai trebuie să menționez că am scris volumul acesta dintr-o suflare – în timp ce mă ocupam, ca de obicei, de o sută de alte lucruri; de la primul cuvânt pus pe hârtie și până la ultimul punct nu a durat mai mult de zece săptămâni. Aș fi vrut să fi avut mai mult timp la dispoziție, și simt și acum că textul nu este destul de rafinat, că ar fi putut să beneficieze de mai multe exemple etc. – am acceptat însă argumentele colegilor mei care spun că, în astfel de situații, când dorești să ai un impact în dezbaterile publice, binele de acum este mai important decât foarte binele de peste un an.

Pe lângă fericirea pe care o resimt fiindcă aduc o contribuție pentru societatea românească, scrierea acestui volum a fost importantă și pentru fericirea mea, căci am redescoperit plăcerea de a scrie în limba română. N-am mai făcut asta de prea mult timp, decât poate doar în texte mici, informale, conversații scrise sau

acte administrative. Un text de o anvergură mai mare, cum este o carte, are o anumită cadență și o anumită solemnitate tipică limbii în care îl scrii, și, da, am redescoperit ceva ce nu știam că am pierdut. O să continui acest exercițiu – cathartic pentru mine, util poate pentru alții –, căci mi s-a clarificat acum și această responsabilitate, spre care oameni mai înțelepți ca mine m-au mai îndemnat în timp, dar pe care acum o văd ca pe un imperativ moral: să scriu în limba mea maternă, chiar dacă discursul științific în lumea profesională din care fac parte se poartă în altă limbă.

1. Introducere în evaluarea standardizată

Ce e un test?

Mă voi referi pe parcursul acestui volum la „teste“ ca fiind proceduri sau metode care au ca obiectiv examinarea sau determinarea prezenței unui factor sau a unui fenomen din sfera diferențelor individuale și care cuprind un set de itemi standardizați (întrebări, stimuli sau sarcini) care sunt scorați într-o manieră standardizată.

Această definiție este cu siguranță corectă, deși poate cam savantă și sforăitoare. Așadar trebuie desfășurată – și voi face asta clarificând câteva noțiuni, patru fiind în mod special importante: „itemi“, „standard“, „factor/fenomen“ și „diferențe individuale“. Pe parcursul acestei cărți voi analiza pe larg problema „itemilor“, așa că nu intru acum în detalii. Voi discuta însă în cele ce urmează despre noțiunile de „standard“ sau „standardizat“, precum și despre „factor/fenomen“ și „diferențe individuale“.

Ce trebuie spus însă de la început – și asta cred că se va vedea pe parcursul întregului volum – este că testele standardizate există în atât de multe forme, prezentări, modalități de administrare, au atât de multe utilități... reprezintă un întreg univers. La nivelul simțului comun sunt confundate cu „testele-grilă“ – în realitate avem de-a face cu un tablou complet diferit: instrumente de măsurare versatile, bazate pe știință serioasă, aplicată cu multă grijă nu doar pentru a oferi măsurare obiectivă, ci și pentru a oferi șanse egale tuturor elevilor testați. Aplicațiile testelor standardizate în școală sunt multiple și generoase, fiind imposibil de ignorat în educația modernă.

Teste vs chestionare

Probabil că acesta este momentul în care trebuie să fac o precizare importantă. De vreme ce subiectul acestei cărți este *testarea* standardizată, mă voi referi la „teste“ ca fiind diferite de „chestionare“ sau „inventare“. În reprezentarea populară ele se confundă, dar în realitate sunt radical diferite. De exemplu, un chestionar de măsurare a atitudinilor, un inventar de măsurare a unor preferințe (e.g., trăsături de personalitate) sau un sondaj de opinie ar putea fi cu ușurință denumite „teste“, fiind de regulă confundate cu acestea. Diferența este că aceste inventare, sondaje, chestionare nu sunt probe stricte, ci culeg răspunsuri care sunt autoportări, de regulă fără a exista un răspuns corect sau greșit.

Să ne imaginăm că într-un chestionar de personalitate apare enunțul sau întrebarea (implicită) „Îmi fac ușor prieteni noi“. Orice răspuns este posibil la această întrebare fără să atragă o penalizare – oamenii sunt diferiți și au preferințe diferite în interacțiunea cu ceilalți. Răspunsul „deloc“ este la fel de legitim ca și „întotdeauna“; ambele răspunsuri reușesc să descrie o preferință, iar preferința introvertită (a elevului mai retras, care nu-și face ușor prieteni noi) nu este în vreun fel mai dezirabilă decât cea extrovertită (a elevului care interacționează facil cu ceilalți și își face ușor prieteni noi). La fel într-un sondaj de opinie în care suntem întrebați dacă preferăm să votăm de regulă conservator sau progresist. Sau într-un chestionar de măsurare a atitudinii față de imigrație în care suntem întrebați dacă acceptăm muncitori imigranți din Africa în comunitatea noastră. Toate acestea măsoară o preferință, o opinie, o atitudine – dar nu o capacitate.

Testele, pe de altă parte, măsoară o capacitate – reflectată de performanța în sarcini care au răspunsuri corecte sau greșite. Până la urmă, atunci când măsurăm o capacitate este posibil să constatăm absența ei sau nivelul ei mai degrabă modest. Itemii unui test, așadar, presupun răspunsuri care sunt corecte sau, dimpotrivă, greșite. Răspunsul la întrebarea „Care din cele două numere este mai mare: 10⁹ sau 9¹⁰?“ nu este în vreun fel o chestiune de opinie, ci presupune capacitatea de a rezolva inegalitatea. Unii

vor răspunde corect, iar alții – greșit. Itemul presupune manifestarea unei performanțe – performanță la care e posibil succesul, la fel cum e posibil și eșecul.

Ce e un test standardizat?

Acum știm ce este un test – dar ce este un test *standardizat*? Adevărul e că ideea de standardizare este parte integrantă din chiar definiția testului – am putea spune, așadar, că „test standardizat“ reprezintă o tautologie: evident că orice test are o componentă de standardizare! Totuși, ce ar trebui să fie caracteristic pentru un test pentru a-și merita cu prisosință acoladele suplimentare de „test standardizat“?

Cea mai des întâlnită confuzie privind testele standardizate este aceea care pune semnul egal între „test standardizat“ și „test-grilă“. Testele standardizate pot fi teste-grilă, dar această caracteristică nu este obligatorie. Cu alte cuvinte, există multe alte feluri de teste standardizate, care nu sunt teste-grilă. Și apoi nu toate testele-grilă sunt teste standardizate.

Conceptul de „standardizare“ în evaluare evocă o caracteristică fundamentală, care este de regulă asociată cu corectitudinea și tratamentul egal – deci cu deziderate sociale, spre care tindem. Anume că toți cei care vor fi evaluați (în acel proces) vor fi tratați la fel.

Mihai și Andrei sunt elevi într-o clasă de a V-a. La finalul anului, ei sunt evaluați printr-un test la biologie. Dacă Mihai primește un test scurt și ușor de 4 întrebări, iar Andrei un test mai lung și mai dificil de 10 întrebări, vom fi cu toții de acord că cei doi copii nu au fost evaluați corect. Vom spune: este normal să primească același test – sau, dacă primesc teste diferite, ne așteptăm ca ele să fie echivalente: să aibă măcar același număr de întrebări și profesorul să facă tot efortul posibil ca întrebările din cele două teste să fie cam la fel de dificile.

Asta nu e însă suficient: evaluarea nu ține doar de test, ea este un întreg proces, care se concentrează și pe felul în care testul este administrat, scorat și interpretat.

Lui Mihai i s-a explicat pe îndelete cum trebuie abordat testul și i s-au dat și câteva exemple. Lui Andrei i s-a cerut să citească singur instrucțiunile. E posibil ca aceste diferențe în felul în care a fost făcut instructajul să-l fi dezavantajat pe unul dintre ei și să fi dus în final la diferențe în scorurile lor.

Mihai a dat testul într-o zi normală de vară, când era răcoare și liniște în clasă și s-a putut concentra. Andrei a dat testul a doua zi, când aparatul de aer condiționat era stricat, căldura era insuportabilă, iar muncitorii care reparau aerul condiționat interpretau *Simfonia a IX-a* la bormașină. Performanța lor e posibil să fi fost diferită din cauza contextului în care au fost testați.

Răspunsurile lui Mihai au fost evaluate prin prisma tuturor pașilor făcuți în încercarea de a rezolva exercițiile, nu doar pe baza rezultatului final. Răspunsurile lui Andrei au fost evaluate doar prin prisma rezultatului final pe care l-a marcat. E posibil ca procesul de scorare să-l fi dezavantajat pe unul dintre ei, ducând la diferențieri și în scorurile lor.

Lui Mihai rezultatul obținut i-a fost explicat în termeni de puncte forte și zone în care mai trebuie să lucreze. Lui Andrei i s-a comunicat doar nota finală. E posibil ca procesul de interpretare și feedback să-l fi dezavantajat pe unul dintre ei și să fi provocat totodată diferențieri în consecințele pe care le are testul.

Și diferențele pot continua... dar, în mod cert, vom spune, pe baza exemplului de mai sus, că Mihai și Andrei nu au fost tratați la fel. Principiul testării standardizate este acela că toți cei testați vor fi tratați identic în toate momentele importante (voi mai discuta despre variațiunile permise, căci există și acestea): același test, același instructaj, aceleași condiții de testare, același mod de scorare etc.

Așadar orice test prin care se fac eforturi pentru ca toate persoanele să fie tratate similar este un test standardizat. Tratatamentul similar privește cel puțin formatul, conținutul, procedura, administrarea și scorarea testului, dar e posibil să vizeze și alte aspecte, mai subtile. Voi discuta mai târziu pe larg despre toate acestea. Până atunci, se impun câteva remarci care țin de fundamentul testării standardizate.