

María Araceli Ruiz-Primo,
professora de la Universitat de Colorado Denver

«No estem acostumats a una cultura d'avaluació que siga formativa»

Purificación Sánchez Delgado. Departament de Mètodes d'Investigació i Diagnòstic en Educació (MIDE). Universitat de València

María Araceli Ruiz-Primo és professora associada a l'Escola d'Educació i Desenvolupament Humà de la Universitat de Colorado Denver, als EUA. El seu treball de recerca es desenvolupa en l'àrea de l'avaluació educativa, tant en l'aprenentatge dels estudiants com en les pràctiques avaluatives i en la formació dels mestres. És directora del Centre de Recerca de l'Escola d'Educació i Desenvolupament Humà i del Laboratori d'Avaluació Educativa, Recerca i Innovació (LEARN, per les sigles en anglès).

En què consisteix la sensibilitat educacional de les proves a gran escala?

La sensibilitat educacional és una característica de les proves que pot captar informació sobre les experiències

d'aprenentatge que té l'alumnat a l'aula. En general, se sap que molt del que mesura una prova a gran escala està relacionat amb el nivell socioeconòmic dels alumnes, és a dir, a major nivell socioeconòmic, major rendiment acadèmic, i a menor nivell, menor rendiment.

Com que les proves són educacionalment no sensibles, no mesuren les oportunitats d'aprenentatge que té l'alumnat a l'aula. Per tant, la variabilitat en els resultats acadèmics és deguda més al nivell socioeconòmic que a la qualitat de la formació que reben els estudiants. El concepte de sensibilitat educacional sorgeix de la necessitat de considerar l'aspecte de qualitat de la formació o oportunitat d'aprenentatge. Si cal prendre decisions sobre el sistema educatiu, que es prenguen, si més no parcialment, sobre la base de la qualitat educativa que rep l'alumne a classe.

Llavors, és obvi que la sensibilitat educacional és un aspecte fonamental en les avaluacions. Però, per què?

Perquè es dona per fet que les puntuacions dels estudiants en les proves a gran escala reflecteixen l'impacte de la formació, però no hi ha cap evidència que aquest siga el cas. Les proves de rendició de comptes (*accountability*) són incapaces de distingir entre educació efectiva, de qualitat, i educació no efectiva i, per tant, no s'haurien d'utilitzar com un indicador vàlid per prendre decisions sobre la qualitat de l'educació. És clar que hi ha molts factors que influeixen sobre les puntuacions dels estudiants en una prova a gran escala, però hauríem de ser capaços de poder adjudicar, almenys parcialment, les puntuacions altes a una formació efectiva i les baixes, a una forma-

Araceli Ruiz-Primo i Purificación Sánchez Delgado.



ció poc efectiva. Això és impossible amb les proves actuals. Fins ara, almenys als Estats Units, el resultat en la majoria de les proves utilitzades per a rendició de comptes és més influenciat per la composició dels estudiants a l'escola i les seues característiques –com ara el nivell socioeconòmic– que pels esforços dels centres escolars per proporcionar una educació efectiva.

I, què us ha portat a interessar-vos en aquesta línia de treball?

Jo treballava a la Universitat de Stanford, al final dels noranta, a l'Stanford Educational Assessment Laboratory (SEAL) i la National Science Foundation (NSF) havia invertit molts diners en projectes del programa *Sistemic Initiatives*. Van ajudar amb milions de dòlars el treball conjunt entre universitats i els districtes escolars en programes de desenvolupament professional de mestres per veure l'impacte que tenien en l'aprenentatge dels estudiants en ciències i matemàtiques. Els resultats no van ser tan positius com es preveïen; l'aprenentatge dels alumnes no havia millorat després del programa. Llavors, NSF va encomanar a SEAL estudiar per què les iniciatives no mostraven els resultats esperats. A partir d'ací es va a analitzar com s'havien fet les avaluacions dels projectes i un aspecte en comú que hi havia entre les avaluacions era que s'utilitzaven com a mesura les proves a gran escala, principalment estatals. En qüestionar-nos per la quantitat d'ítems en una prova estatal que poden cobrir el que han estudiat els alumnes a classe, sorgeix la idea de distància dels ítems al currículum. És important recordar que als EUA no hi ha un currículum nacional.

Quines repercussions pot tenir la sensibilitat educacional en les avaluacions a gran escala?

Actualment, als EUA s'està proposant la sensibilitat educacional com una característica tècnica de les proves a gran escala. No passarà molt de temps perquè les proves estatals i nacionals hagen de donar compte d'aquesta sensibilitat.

Quines recomanacions faríeu per considerar la sensibilitat educacional en les avaluacions a gran escala?

Dissenyar estudis que puguin recollir l'evidència empírica que les proves són sensibles a la qualitat formativa per poder prendre decisions més informades. Tenir en compte la sensibilitat educacional i proveir d'ítems diferents distàncies. Sempre he pensat que no es poden prendre decisions basades en una sola font. Si s'ha d'avaluar el mestre, si el van a acomiadar, o es va a tancar una escola, no ens podem basar en una única prova. En un plantejament utòpic seria excel·lent un sistema d'avaluació amb diferents fonts d'informació, almenys tres: els mestres, perquè són qui millor coneix l'alumnat; la informació de cada districte amb les seues proves específiques; i les proves a gran escala, estatals, nacionals o internacionals que s'utilitzen per a rendició de comptes.

I l'avaluació de l'ensenyament per competències?

En les nostres recerques i en altres s'ha trobat que la majoria dels ítems de les proves a gran escala, no menys del 65-70%, són del tipus declaratiu, pocs ítems mesuren la comprensió profunda dels conceptes; el que anomenem coneixement de tipus esquemàtic. Un enfocament en competències ha d'avaluar més enllà del conceptual; necessàriament requereix repensar què s'està fent a les aules per promoure competències. I això ens porta als contextos, escenaris dels ítems... No hi ha una teoria del desenvolupament d'ítems contextualitzats. Qui diu que els ítems que tenen un context o escenari són sempre millors i mesuren competències?

Nosaltres hem plantejat la importància de saber quin tipus d'ítem es necessita per avaluar competències. Des d'aquesta perspectiva, cal avaluar un coneixement que siga més estratègic, que és cap a on van les competències. Si bé això és difícil, és cert que és factible.

Hi ha investigadors als EUA que estan desenvolupant projectes d'avaluació per ordinador sobre la base d'ítems enfocats a la solució de problemes, on es valora no solament la resposta de l'ítem sinó el procés i les estratègies que s'han utilitzat per resoldre'l. Caldria esperar que eixes estratègies no siguen preses a l'atzar sinó basades en un coneixement expert. Això va lligat a un altre tema que m'apassiona, l'avaluació formativa: quan a l'estudiant l'ensenyes a desenvolupar estratègies metacognitives, l'ajudes a desenvolupar una de les anomenades *competències del segle XXI*. Una manera de fer-ho és mitjançant l'avaluació formativa, que és molt més que recolzar la metacognició.

A més, sabem per TIMSS (Trends in International Mathematics and Science Study) i PISA (Program for International Student Assessment) que els països amb les puntuacions més altes ensenyen als estudiants amb major profunditat. No pensen a donar molts continguts sinó a treballar-ne pocs però a fons. Cal una cultura de l'avaluació i, ser més assertius. Dir les coses pel seu nom. Tenim por de dir que això no funciona i prendre responsabilitats.

Finalment, com relacionaríeu la sensibilitat educacional amb el currículum educatiu, amb les metodologies didàctiques utilitzades a les aules i amb el tipus de procediment avaluatiu que utilitzen els mestres?

Sí, tenen molt a veure. Des d'una perspectiva a gran escala, molt del que s'ha tractat per avaluar la sensibilitat educacional d'una prova, s'ha fet com a mínim si l'ítem està apuntant a alguna cosa que mínimament es va veure a l'aula. Això s'ha avaluat, per exemple, amb qüestionaris als mestres (o als alumnes). A un altre nivell més profund es pot estudiar de prop la qualitat de les oportunitats d'aprenentatge que es van donar als estudiants. Si un ítem no pot ajudar a determinar la qualitat de la formació que s'ha rebut, llavors com podem prendre decisions sobre el sistema educatiu en què està l'estudiant? ■