

AVALUACIÓ I DEMOCRÀCIA EN ELS SISTEMES EDUCATIUS

Georges Felouzis. *Professor de Sociologia de la Universitat de Ginebra.*

Georges.Felouzis@pse.unige.ch

Avui en dia, l'avaluació és una eina clàssica de posada en qüestió de les polítiques públiques. Es tracta d'un *instrument* a disposició de les polítiques públiques per rendir comptes sobre la seva acció i per orientar la seva aplicació (Lascoume i Le Galès 2004), però també, i potser sobretot, es tracta d'una eina de coneixement de la qual poden apropiarse els actors per tal de comprendre les conseqüències de la seva activitat quotidiana, d'allò que d'alguna manera "produeixen". Aquest és un punt central que pren una dimensió particular en el món escolar. Com ho demostren la majoria de treballs, els fenòmens escolars són d'una gran complexitat i comprendre les conseqüències d'una acció (en una aula, un centre educatiu o, a nivell més "macro", una regió o una nació) només és possible si es fa un càlcul objectiu dels fenòmens estudiats. Així mateix, l'avaluació —en tot cas el concepte que defensaré aquí— no és altra cosa que un càlcul objectiu de les conseqüències de l'acció pública, sigui quin en sigui el nivell (individual o col·lectiu).

Aquesta definició prèvia no implica en absolut que aquesta avaluació sigui "neutra" ni el fruit pur del coneixement objectiu. Tot i que es desprengui d'aquest coneixement, sempre és fruit d'una interpretació i de les idees preconcebudes que la regeixen. L'avaluació només es concep a partir dels objectius i de la finalitat que es desprenen de les opcions polítiques i socials, de les preferències socials valoritzades per la societat en un moment donat. Per tant, ens hem de demanar alhora la *fiabilitat* tècnica de la mesura, però també les *condicions de producció* i la *utilització política* que se'n fa. Defensaré aquí la idea que l'avaluació en l'àmbit de les polítiques escolars pot ser un mitjà per accedir a un plus de democràcia, però que aquesta capacitat democràtica és només una possibilitat entre d'altres. Òbviament, no és intrínseca a l'avaluació mateixa. Segons l'ús que se'n faci, pot convertir-se en un mitjà de dominació política (entre nacions desiguals, per exemple), de dominació social (entre grups socials), i fins i tot de reproducció i conservació de les desigualtats.

Segons l'ús que se'n faci, l'avaluació pot esdevenir un mitjà de dominació política o social

Les polítiques públiques en matèria d'educació, com en d'altres àmbits de la vida social, es conceben avui a partir de les eines d'avaluació. Des del moment que l'estat ja no és només *dirigista* sinó que esdevé *regulador*, ha d'avaluar, "contractualitzar" i fixar objectius als actors en un marc preestablert. Aquesta nova concepció de les relacions entre l'Estat i la societat té com a conseqüència natural la transformació de l'estatus de l'avaluació, que s'ha convertit en una part integrant de les polítiques públiques. Ja no és un element que s'hi afegiria *a posteriori*, sinó una dimensió determinant de l'acció política en ella mateixa. Allò que passa a França i en altres països europeus en matèria d'ensenyament superior (contractualització de les universitats, autonomia dels centres, etc.) i d'ensenyament secundari (autonomia dels centres regulada per l'estat a través dels programes i dels currículums): qualsevol política ha d'integrar els criteris de la pròpia avaluació.

En aquest marc ràpidament esbossat, què podem dir sobre les relacions entre avaluació i democràcia? La meva idea és que aquesta relació ja és antiga. Si considerem moltes de les produccions de ciències socials en l'àmbit escolar a partir dels anys 1960 als Estats Units i a Europa, podem considerar-les majoritàriament com a avaluacions de les polítiques escolars. Les més notables, a més, han contribuït a democratitzar els sistemes educatius dels països que analitzaven. Estic pensant particularment en els treballs de James Coleman (1966) i de Christopher Jencks (1978) sobre les desigualtats escolars entre els estudiants blancs i negres. Mostrar les conseqüències socials i econòmiques d'aquestes desigualtats, i mostrar en quins aspectes són fruit del sistema educatiu, és un exemple idealtípic d'allò que és —i hauria de ser— l'avaluació en matèria escolar: partir de preguntes fortes, proposar mesures múltiples i objectives i, així, crear les condicions d'una democratització dels sistemes educatius estudiats. Em vénen al cap espontàniament altres exemples. Els treballs de Pierre Bourdieu (1964) i Jean-Claude Passeron (1970) sobre el sistema escolar francès, que mostraven el paper de l'escola en el procés de reproducció de la dominació social i econòmica. No és aquest "descobriment" —un dels més debatuts sobre el sistema escolar francès— un model d'avaluació possible? Cal subratllar que aquests treballs han participat àmpliament en el moviment de democratització del sistema educatiu francès durant els anys 1960 i 1970. Podria citar d'altres exemples. Penso en els treballs de Basil Bernstein (1975) sobre la distància lingüística i social que es troba en els fonaments de les desigualtats escolars en el sistema educatiu anglès: aquí també la sociologia s'uneix a l'avaluació per posar "a prova els fets", les ideologies escolars, les pràctiques pedagògiques i les polítiques educatives.

Si acceptem definir la democratització escolar com una *reducció de desigualtats* entre individus i entre grups, podem avançar que aquests treballs dels sociòlegs han contribuït àmpliament a democratitzar els sistemes d'ensenyament objecte d'estudi. Vet aquí un element determinant per reflexionar sobre la relació entre avaluació i democràcia, ja que aquests treballs dels sociòlegs només han pogut participar en la democratització dels sistemes educatius perquè han estat difosos socialment i han participat en els debats públics. La meua tesi és, per tant, que l'avaluació només té capacitat democràtica si, i només si, es difonen els seus resultats, i així participen en el debat públic. Si els actors se n'apropien, bé com a mitjà per millorar les seves pràctiques (en l'àmbit de centres educatius, per exemple), bé com a argument utilitzat en el debat públic en un moment donat, la relació entre avaluació i democràcia no és, doncs, directa: sempre està mediatitzada per la seva apropiació per part dels actors en el context d'un debat públic democràtic. Posaré dos exemples per il·lustrar la meua tesi. El primer té a veure amb les comparacions internacionals. Es basen en l'ús dels resultats de PISA 2000 i 2003 en un cert nombre de països europeus. Agafaré l'exemple d'Alemanya, França i Suïssa. El segon exemple tracta més sobre el nivell dels centres educatius. Mostraré com, en el cas francès, aquestes avaluacions poden tenir efectes perversos i accentuar les desigualtats socials i escolars entre l'alumnat i els centres educatius.

L'avaluació només té capacitat democràtica si, i només si, es difonen els seus resultats i així participen en el debat públic

El primer té a veure amb les comparacions internacionals. Es basen en l'ús dels resultats de PISA 2000 i 2003 en un cert nombre de països europeus. Agafaré l'exemple d'Alemanya, França i Suïssa. El segon exemple tracta més sobre el nivell dels centres educatius. Mostraré com, en el cas francès, aquestes avaluacions poden tenir efectes perversos i accentuar les desigualtats socials i escolars entre l'alumnat i els centres educatius.

Els estudis PISA i les conseqüències en el debat escolar en tres països europeus

La reeixida aposta dels estudis internacionals duts a terme per l'OCDE consisteix a comparar sistemes educatius amb una estructura, una durada i uns objectius que poden ser molt diferents. Com comparar sistemes educatius que no tenen els mateixos programes, els mateixos mètodes ni la mateixa duració temporal? La solució emprada per aquests estudis consisteix a comparar no els sistemes entre ells sinó les *conseqüències* d'aquests sistemes en termes d'adquisició de competències i coneixements (comprensió lectora, cultura matemàtica i científica), de desigualtats, d'efectes produïts pels centres, de relacions entre professorat i alumnat, etc.

Tot l'interès d'aquests treballs resideix òbviament en el seu rigor metodològic i en la seva capacitat de comparar i, per tant, d'*avaluar* els sistemes educatius: cada país se situa segons la puntuació obtinguda pels joves de 15 anys en tal o tal escala (comprensió lectora, cultura matemàtica, etc.). I els resultats tenen molt d'interès per avaluar *l'eficàcia*, *l'equitat* i *el rendiment* de cada sistema. Però és evident que això no és tot.

L'avaluació dels sistemes educatius no és només una qüestió de "notes", sinó també de l'anàlisi dels factors de desigualtat que condicionen els aprenentatges. Cal calcular els "efectes de context" (en què depèn dels centres el rendiment?), el pes dels factors socioeconòmics, la influència d'un estatus de migrant o d'al·lòcton, etc. El càlcul de l'efecte d'aquestes circumstàncies sobre les competències dels joves de 15 anys és una avaluació molt interessant, ja que permet una comparació entre països. Les diferències i les desigualtats ja no es veuen "neutralitzades", sinó tot al contrari, relativitzades: si el fet de ser migrant o al·lòcton és un obstacle molt fort aquí i molt feble allà, no serà que els sistemes educatius i els modes d'integració dels migrants i dels seus descendents marquen aquí la diferència? Si el nivell sociocultural té un efecte variable segons les nacions, no serà perquè cada sistema educatiu proposa concepcions educatives jerarquizades segons criteris contrastats i, per tant, necessàriament relatius?

Així, doncs, què aprenem d'aquests estudis? Doncs que:

- els països amb puntuacions mitjanes més altes tenen també una dispersió feble (són eficaços i justos), mentre que, al contrari, les desigualtats són més fortes en els països on la puntuació mitjana és feble;
- la riquesa i la inversió en educació expliquen una part d'aquests resultats, però només una part. Les explicacions també s'han de buscar en l'organització dels sistemes, en el funcionament dels centres i en d'altres factors que encara queden per descobrir. No és anodí, per exemple, veure que les desigualtats socioeconòmiques poden corregir-se, en certa mesura, en l'àmbit de centres educatius (PISA 2003);
- els sistemes democratitzats (en el sentit que proposen durant el màxim temps possible el mateix currículum), són globalment més eficients que els altres;
- en definitiva, podem veure tot el *pes de la política* en la definició mateixa de les desigualtats i dels resultats de l'alumnat. Des del moment que el nivell de competència de l'alumnat i el nivell de les desigualtats en una societat donada depenen de factors controlables per les polítiques educatives, els estudis PISA constitueixen una avaluació d'aquestes polítiques.

L'avaluació dels sistemes educatius no és només una qüestió de "notes". Es tracta també d'analitzar els factors de desigualtat que condicionen els aprenentatges

Em proposo analitzar la recepció d'aquests estudis en tres països per intentar comprendre quines són les relacions entre aquests estudis avaluatius i la democratització dels sistemes educatius en aquests països.

■ **A França**, la recepció de l'estudi PISA suscita sempre debats animats. No només perquè l'ensenyament té un lloc important en l'imaginari nacional, sinó també perquè les comparacions internacionals es viuen massa sovint com una posada en qüestió dels principis del sistema escolar francès. I, no obstant això, la posició de França en les comparacions internacionals no és deshonorant: se situa en la mitjana de l'OCDE pel que fa a la comprensió lectora, i més aviat en bona posició pel que fa a la cultura científica i matemàtica. La inversió educativa se situa també en la mitjana dels països. Però el problema francès en matèria d'educació és doble. Hi ha, per començar, una reticència forta a ser comparat amb altres països, i a més les desigualtats són encara grans (desigualtats socials, culturals i vinculades amb el context dels centres). Per tant, el debat a França s'orienta en dues direccions:

- La primera ha estat la de criticar els estudis PISA per considerar-los massa marcats pels models educatius i culturals anglosaxons. Els pressupòsits i els implícits de l'estudi serien aliens (fins i tot oposats) a la cultura escolar francesa, cosa que explicaria els resultats mitjans dels joves de 15 anys a França. Aquesta argumentació no és il·legítima en el sentit que sempre és necessari qüestionar les condicions de producció dels estudis comparatius tan genèrics i "avaluar l'avaluació". Altres crítiques, més basades en l'estudi i la comparació, parlen dels límits d'aquestes comparacions. Són les idees desenvolupades per E. Bautier, J. Crinon, P. Rayou i J-Y Rochex (2007). Proposen una anàlisi crítica del concepte i del càlcul de la comprensió lectora demostrant que, de fet, els estudis PISA calculen competències molt heterogènies, ja que "les maneres de fer depenen i varien segons els contextos" (p. 98). Per tant, cal estar molt atent a allò que es calcula per extreure totes les conseqüències dels resultats publicats.
- La segona direcció ha estat un veritable debat sobre les qüestions escolars: la repetició de l'alumnat en l'ensenyament primari, els mètodes d'aprenentatge de la lectura al principi de l'escolaritat, el tema del sistema educatiu i de l'escola secundària única, la carta escolar (que obliga a matricular els fills o filles al centre públic proper al domicili familiar) i la segregació social i ètnica en els centres educatius, l'horari i la feina del

A França,
l'estudi PISA ha
provocat un veritable
debat en diversos àmbits
de les polítiques
educatives

professorat. Tots aquests debats sobre l'escola no tenen com a font única les avaluacions de PISA, però n'han estat alimentats en el sentit que aquestes avaluacions proposen una mirada comparativa "externa" que relativitza les tries nacionals.

■ **A Alemanya**, la situació és molt diferent i les opcions del sistema educatiu alemany són oposades a les de França. L'ensenyament proposa una orientació molt precoç i selectiva a partir del final de l'ensenyament elemental. A partir dels 10 o 11 anys, s'orienta l'alumnat bé cap a l'ensenyament general que porta als estudis superiors (el *Gymnasium*), bé cap a la *Realschule*,¹ o bé cap a la *Hauptschule*.² Així, amb 15 anys, molts joves no han seguit l'ensenyament general que els havia de permetre adquirir les competències necessàries. Això produeix desigualtats particularment fortes, les més importants dels països que han participat en el PISA 2000. Hem d'afegir que aquestes desigualtats de competències retallen les característiques dels joves: els immigrants i joves que surten de la immigració, així com aquells que provenen d'un mitjà socioeconòmic desafavorit, estan particularment desavantatjats. Aquests resultats han estat viscuts a Alemanya com un veritable xoc (Ertl 2006), que ha suscitat un debat i una reflexió sobre el sistema educatiu (Koch 2006). L'orientació precoç de l'alumnat no

A Alemanya, els resultats en el PISA 2000 van produir un veritable xoc que va suscitar debat i reflexió sobre el sistema educatiu

s'ha posat en dubte. El debat s'ha centrat en els estàndards educatius nacionals i la posada en marxa de programes ràpidament aplicats. Aquesta reforma, impensable sense els resultats de l'estudi PISA, també ha permès una reflexió sobre els ritmes escolars i l'educació dels joves que provenen de la immigració. De tota manera, subratllem amb H. Ertl que les conseqüències d'aquestes reformes estan lluny d'estar assegurades si tenim en compte les prerrogatives dels *Länder* en matèria educativa i la forta autonomia del professorat i dels centres educatius.

■ **A Suïssa**, els resultats de PISA 2000 estan molt per sobre de la mitjana dels països de l'OCDE pel que fa a les matemàtiques i a la cultura científica. La comprensió lectora, però, es troba just en la mitjana, i fins i tot una mica per sota (Nidegger 2005). A més, les puntuacions mitjanes amaguen contrastos molt forts entre els cantons. Per això, Suïssa va decidir fer les proves a més alumnes dels previstos per introduir

¹ Sis anys d'escolarització que permeten accedir a qualsevol formació professional o a un batxillerat especialitzat.

² Cinc o sis anys d'escolarització (segons els Länder) que permeten accedir només a formació professional.

la variable "cantó" en les anàlisis de PISA 2003 (PISA 2005). Els resultats tenen gran interès, ja que mostren les disparitats de competències entre els joves de 15 anys segons la regió i el cantó. Per exemple, la puntuació mitjana en matemàtiques és de 542 en la Suïssa alemanya, de 528 en la Suïssa francòfona i de 511 en la Suïssa italiana. Si considerem els cantons, observem una puntuació molt alta per a Friburg (553) i per a Saint-Gal i Turgòvia (551). Les puntuacions més febles són per a Vaud (524), Ticino (510) i Ginebra (508). L'explicació d'aquestes disparitats no s'ha dilucidat completament (diferenciació dels *curricula*, nivell socioeconòmic dels cantons, etc.). A més, l'estudi de Moreau, Nidegger i Soussi (2007) mostra que les categories utilitzades per a les comparacions internacionals a gran escala no són sempre adients en l'àmbit local. És el cas de l'àmbit socioeconòmic: "la relació entre els resultats i el nivell socioeconòmic sembla ben real, però l'índex emprat per PISA no és pertinent per a certes comparacions entre cantons" (p. 47). Quines han estat les conseqüències de PISA en el sistema educatiu suís? En una perspectiva propera a la d'Alemanya, s'ha creat un projecte d'harmonització (HarmoS). Es tracta de definir els estàndards i els models de competència comuns als vint-i-sis cantons. L'harmonització està en procés de desenvolupament.

A Suïssa s'ha creat un projecte d'harmonització del sistema educatiu. És una conseqüència dels resultats de l'estudi PISA

El paper de l'avaluació, com es pot veure, és el de posar a prova els resultats dels sistemes educatius avaluats. És una manera de *regular* els sistemes educatius en comparar els seus resultats amb els objectius propis, així com amb els altres sistemes educatius del món. Però aquesta regulació no s'opera de forma immediata, ja que els efectes de l'avaluació, especialment les conseqüències en l'eficàcia i l'equitat dels sistemes educatius, no són mai el fruit d'una relació directa entre l'avaluació, d'una banda, i les reformes, de l'altra. És amb l'apropiació dels resultats de l'avaluació per part dels actors polítics i dels ciutadans que l'avaluació esdevé un factor de reforma i d'evolució dels sistemes educatius.

És amb l'apropiació dels resultats de l'avaluació per part dels actors polítics i dels ciutadans que l'avaluació esdevé un factor de reforma i d'evolució dels sistemes educatius

Fins ara, només he tractat la qüestió de l'avaluació a partir de l'estudi PISA. Per comprendre el paper democràtic de l'avaluació, és necessari observar més en detall les condicions de producció i els usos de les avaluacions en l'àmbit escolar. Per fer-ho, recordaré el cas francès, en primer lloc perquè és el que conec millor, però també perquè es tracta d'una nació on les eines d'avaluació de les polítiques escolars estan molt desenvolupades. És, crec, un dels avantatges

d'un estat fort i centralitzat. Aquest "zoom" damunt el cas francès em permetrà concretar en quines condicions els resultats de l'avaluació poden tenir virtuts democràtiques, és a dir, de millora dels sistemes educatius en el sentit d'una major eficàcia i d'una major equitat.

Allò que produeix l'avaluació dels centres: el cas dels indicadors de resultats dels instituts (*lycées*) de França

Em proposo tractar el tema de l'avaluació a partir d'un estudi de cas que vaig publicar en un article de 2005 a la *Revue Française de Sociologie*. Es tracta d'una anàlisi crítica dels indicadors de resultats dels instituts. Els indicadors ens interessen per més d'una raó. En primer lloc, es tracta d'eines d'avaluació posades en marxa per la DEP (Direcció de l'Avaluació i de la Prospectiva del Ministeri d'Educació) per oferir informacions fiables i reproduïbles sobre la qualitat dels instituts. Aquests indicadors han estat concebuts per la DEP com a informacions per al públic en general que es difonen mitjançant la premsa i Internet. També es tracta "d'eines de pilotatge", és a dir, que tenen la vocació de ser utilitzades pels equips de professorat dels instituts per tal de millorar els resultats (si no estan a l'alçada de les expectatives) o mantenir-los. Per tant, ens trobem amb un cas típic d'avaluació, la vocació de la qual és millorar l'eficàcia dels centres i així limitar al màxim possible les desigualtats que es produeixen per l'heterogeneïtat de l'oferta escolar.

El context de creació d'aquests indicadors és interessant. Cap a finals dels anys 1980, una part de la premsa publicava un palmarès dels instituts que només prenia en consideració uns indicadors molt generals sobre l'eficàcia dels centres; tenia en compte, senzillament, el percentatge d'èxit brut en el batxillerat. Alguns centres tenien el 99% d'èxit i d'altres amb prou feines arribaven al 40 o 50%. Això vol dir necessàriament que els primers eren bons instituts i els altres, molt dolents? Avui sabem que això no és així, ja que cal tenir en compte la situació de partida dels centres. Alguns tenen alumnes amb un baix rendiment des del començament, d'altres tenen molt bons alumnes des del començament. Per tant, era important produir una comparació justa. Per aquest motiu, el Ministeri d'Educació va posar en marxa aquests indicadors per comparar els centres educatius amb bases sòlides. La idea d'aquesta comparació és la de neutralitzar les diferències de partida degudes a la naturalesa del públic dels centres, essencialment al voltant de dues característiques de l'alumnat: l'edat i l'origen social.

Els indicadors de resultats dels instituts francesos tenen en compte l'edat i l'origen social de l'alumnat

Per establir aquesta comparació, el mètode consisteix a calcular *allò que s'espera* d'un centre educatiu, és a dir, el percentatge d'èxit de l'alumnat que s'hauria de donar tenint en compte l'edat i l'origen social del públic d'aquell mateix centre. És un càlcul relativament senzill que es basa en la hipòtesi que cada centre funciona com la mitjana (bé a França o bé a l'*académie*, la circumscripció administrativa d'ensenyament) del conjunt dels altres centres. Tenint en compte la repartició segons l'edat (avançat, a temps, amb un any de retard i amb dos anys de retard) i l'origen social (molt afavorit, afavorit, mitjà i desafavorit), podem calcular allò que s'espera de cada centre educatiu.

Aquesta expectativa és molt important, ja que es pot comparar amb *allò observat*, que és la realitat del percentatge d'èxit del centre. La comparació d'ambdós indicadors (allò esperat i allò observat) dóna el *valor afegit* de l'institut en qüestió. Es calculen tres indicadors:

- *El percentatge d'èxit en el batxillerat.* És la proporció d'alumnes del curs terminal (el tercer i últim de l'institut) que aproven el batxillerat.
- *El percentatge d'accés al batxillerat.* És la proporció d'alumnes que entren a segon curs (el primer any de l'institut) i que aproven el batxillerat tres o quatre anys després (poden repetir un any).
- *El percentatge de batxillers entre aquells que surten de l'institut.* És la proporció d'alumnes que deixen l'institut i que obtenen el títol de batxiller. Alguns estudiants poden abandonar els estudis, o reorientar-los o sortir a l'estranger, etc. Per tant, no tot l'alumnat que surt de l'institut obté el títol de batxiller.

Per exemple, un institut on s'espera un 75% d'èxit en el batxillerat i els alumnes del qual el passen en un 85% tindrà un valor afegit d'un 10%. Al contrari, un institut on s'espera un 80% d'èxit i només aproven un 70% dels alumnes tindrà un valor afegit negatiu (-10%).

Per tant, les coses semblen força senzilles. Cada institut té un públic particular definit segons l'edat i l'origen social de l'alumnat. La DEP fa la hipòtesi que aquestes dues variables són prou fiables per "predir" l'èxit. En neutralitzar-les, pretén calcular "l'efecte del centre", és a dir, el marge de maniobra del centre i la seva acció pròpia per aconseguir l'èxit de l'alumnat. Tanmateix —i aquells que coneixen els estudis PISA ho saben—, amb l'edat i l'origen social no n'hi ha prou per

Els indicadors de resultats dels instituts francesos pretenen calcular el valor afegit de cada centre educatiu

predir l'èxit. Són variables massa poc matisades per mostrar les oportunitats *a priori* d'èxit escolar. L'origen social, per exemple, és una variable que es distribueix només en quatre modalitats (molt afavorit, afavorit, mitjà i desafavorit), i això només reflecteix de forma imperfecta la realitat social i cultural de l'alumnat. A més, l'edat de l'alumnat quan entra a l'institut (l'edat "normal" són 15 anys) depèn de les polítiques de repetició. Aquestes polítiques van en el sentit de menys repetició a l'escola secundària. Per tant, aquesta variable és probablement imperfecta per descriure el "nivell escolar" efectiu de l'alumnat.

Ens cal, no obstant això, intentar calcular els límits d'aquests indicadors que, he de subratllar-ho, continuen essent publicats per la DEP tot i la demostració que he pogut fer del seu caràcter aproximatiu. Això ens portarà, conseqüentment, a una reflexió més general sobre els usos polítics i socials de l'avaluació.

Com "avaluar l'avaluació"? Si l'edat i l'origen social no són predictors prou fiables del nivell de sortida de l'alumnat, caldrà buscar-ne un altre per tal d'establir la comparació. Aquest predictor s'ha trobat mitjançant un examen que tot l'alumnat passa al final de la seva escolaritat a l'escola secundària, just abans d'entrar a l'institut. Es tracta de l'examen *brevet des collèges*, una part de l'avaluació del qual es fa amb exàmens escrits: tot l'alumnat fa les mateixes proves, en les mateixes condicions i se'l puntua segons els mateixos criteris. Aquesta mesura, és clar, és imperfecta, però l'hem utilitzat per mesurar el nivell general de l'alumnat que entra al primer curs de batxillerat en cada centre. Aquesta mesura, comparada amb els resultats dels indicadors de rendiment, permet d'avaluar l'avaluació: els indicadors permeten realment neutralitzar els efectes de composició dels centres? El valor afegit és realment independent del nivell de sortida de l'alumnat en cada centre?

Per respondre aquestes preguntes, presento la **taula núm. 1**, on s'exposa la nota mitjana de l'alumnat de nou centres (percentatge esperat feble: 62%) amb el valor afegit del centre. Si els indicadors són fiables, no hi hauria d'haver relació entre aquestes dues sèries de dades. Si els indicadors no són fiables, hauria d'aparèixer-hi una relació.

Taula 1 ■ Nivell escolar de l'alumnat de primer de batxillerat l'any 2000 segons el percentatge d'accés esperat per al 2002 (62%) (França)

Percentatge d'accés esperat baix: 62%

Institut	Nota mitjana de l'alumnat en l'examen escrit del <i>brevet</i>	Valor afegit del centre
Institut 1	10,02	-7
Institut 2	10,47	-14
Institut 3	10,56	-1
Institut 4	10,98	-16
Institut 5	11,05	-4
Institut 6	11,15	6
Institut 7	11,34	1
Institut 8	11,50	4
Institut 9	11,67	3

La taula s'ha de llegir així: la nota mitjana a l'examen escrit del *brevet* de l'alumnat inscrit al primer curs en l'institut 1 és de 10,02 sobre 20 i el "valor afegit" d'aquest establiment és de -7. Això significa que el percentatge d'accés del primer curs al tercer (terminal) d'aquest centre per al 2002 és 62 (l'esperat) $-7 = 55\%$.

Observem en aquesta taula que els centres el valor afegit dels quals és negatiu tenien alumnes amb mals resultats des del principi (entre 10 i 11 sobre 20). Aquells que tenen un valor afegit positiu tenien des del principi alumnes escolarment més forts (més d'11 sobre 20). Aquestes diferències de nivell són molt sorprenents ja que s'esperen els mateixos *percentatges d'accés* d'aquests centres. Però no tenen els mateixos alumnes des del principi. Això planteja un problema, ja que se suposa que els indicadors de resultats neutralitzen els factors d'èxit externs i aquí tenim un factor d'èxit extern que és bastant fort: en l'institut 1, l'alumnat té un 10 de mitjana i en l'institut 9, un 11,67. És una diferència que pot semblar insignificant (1,65), però sobre la mitjana és una diferència molt important. Això vol dir que els centres tenen alumnes molt diferents.

Mirem ara el valor afegit dels centres (segona columna). En els centres que reben al principi els estudiants més febles, tenim els valors afegits negatius. Això vol dir que, segons els indicadors de resultats, no compleixen el seu "contracte". D'altra banda, aquells centres que són al final de la taula tenen un valor afegit positiu.

Què ens ensenya aquesta taula?

En primer lloc, els indicadors de resultats dels instituts, *a priori*, no neutralitzen —com se suposa que ho haurien de fer— els factors d'èxit extern; és a dir, que esperen el mateix de centres que tenen alumnes molt diferents. En segon lloc, sembla (ens trobem aquí en el nivell de les primeres impressions) que el valor afegit depèn en part almenys del nivell real de sortida de l'alumnat al primer curs de batxillerat: els instituts amb un valor afegit positiu són també aquells que tenen els millors alumnes al primer curs.

Els centres amb un valor afegit positiu són també aquells que tenen el millor alumnat en arribar a l'institut

Podem dir el mateix dels establiments dels quals s'espera un nivell alt? Es tracta del mateix fenomen?

Taula 2 ■ Nivell escolar de l'alumnat de primer de batxillerat segons el percentatge d'accés esperat per al 2002 (71%) (França)

Percentatge d'accés esperat alt: 71%

Institut	Nota mitjana de l'alumnat en l'examen escrit del <i>brevet</i>	Valor afegit del centre
Institut 1	10,38	1
Institut 2	11,26	-6
Institut 3	11,35	-1
Institut 4	11,51	6
Institut 5	11,53	6
Institut 6	11,65	2
Institut 7	11,73	4

Els resultats de la **taula núm. 2** són similars als de la **taula núm. 1**: es dibuixa una relació entre el nivell real de l'alumnat quan accedeix al primer curs i el valor afegit de l'institut. Si el resultat es confirma, llavors haurem demostrat que els indicadors de resultats dels instituts no aconsegueixen el seu objectiu. Aquests indicadors pretenen proposar una mesura de l'efecte del centre educatiu quan, de fet, només mesuren un efecte de composició, que no és ni de bon tros el mateix. Què podem dir per anar una mica més lluny? L'estudi s'ha centrat en 103 instituts (el conjunt d'una circumscripció administrativa d'ensenyament) i ha consistit a comparar d'una forma sistemàtica el nivell de partida de l'alumnat i el valor afegit dels instituts. Els resultats són molt interessants per comprendre alhora les capacitats i els límits de l'avaluació. Em limitaré aquí a donar els resultats principals (Felouzis 2005).

1. Els indicadors de resultats dels instituts no són fiables. Presenten imperfeccions en la mesura mateixa del nivell de partida de l'alumnat. Per tant, el "valor afegit" calculat només reflecteix de manera imperfecta la capacitat dels centres educatius per fer progressar l'alumnat.
2. Existeix un "biaix" sistemàtic en el càlcul dels indicadors. Els centres que escolaritzen l'alumnat de les classes més populars tenen sempre alumnes més febles d'allò que s'esperava. Els centres més benestants sovint tenen alumnes escolarment més forts d'allò que s'esperava. Això té com a conseqüència que el càlcul del valor afegit és més favorable als instituts benestants que als instituts populars. Els primers arriben sovint a obtenir un valor afegit positiu; els segons tenen problemes per arribar a aquest objectiu. I aquest és el resultat de la manera de calcular els indicadors i no de la "qualitat" real dels instituts.
3. Per què és així? Per què els instituts més benestants tenen alumnes millors d'allò que es podria pensar per la composició social i l'edat del seu públic? I per què és a la inversa per als instituts més desfavorits? Doncs per qüestió dels "mercats escolars": els centres més benestants atrauen els millors alumnes; els altres instituts, els perden.

Els indicadors de resultats dels instituts francesos no són fiables. Presenten imperfeccions en la mesura mateixa del nivell de partida de l'alumnat

Taula 3 ■ Característiques i resultats dels centres segons el nivell de l'alumnat a primer de batxillerat (França)

	Distribució dels instituts segons si els alumnes que reben són més febles o més forts d'allò que s'esperava			
	1r quartil (alumnes més febles del que s'esperava) 26 instituts	2n quartil 25 instituts	3r quartil 26 instituts	4t quartil (alumnes més forts del que s'esperava) 26 instituts
Nombre de centres privats	12	11	6	7
Nombre d'instituts amb més d'un terç d'alumnes desfavorits	15	15	13	6
Nombre d'instituts amb un valor afegit positiu (accés del primer curs al batxillerat)	5	5	16	19
Nombre d'instituts amb un valor afegit positiu (curs terminal/ batxillerat)	6	11	17	18

A causa d'això, els mercats escolars marquen diferències. I el problema dels indicadors de rendiment és que no arriben a tenir en compte aquest fenomen, com il·lustra la **taula núm. 3**. En els instituts del quart quartil és on són més rars els establiments més populars. Però no n'hi manquen, ja que representen gairebé un quart dels centres d'aquesta categoria. Els instituts atractius són sovint els més "burgesos", però no és el cas sistemàticament. Cal notar, de tota manera, que dels sis instituts del quart quartil, on més d'un terç de l'alumnat prové d'un entorn desafavorit, tres són instituts privats, dos dels quals escolaritzen el 44% i 50% dels alumnes desafavorits.

La **taula núm. 3** dona el nombre d'establiments de cada quartil que produeix un valor afegit positiu, tant respecte del percentatge d'accés com del percentatge d'èxit en el batxillerat. Aquests resultats ens porten a fer dos comentaris. El primer és que els instituts del primer quartil —aquells que reben alumnes més febles d'allò que s'esperava— rarament aconsegueixen produir un "valor afegit". I, inversament, els instituts que reben alumnes més forts d'allò que s'esperava —quart quartil— ho aconsegueixen molt sovint. Això es verifica pels dos indicadors mantinguts: el percentatge d'accés i el percentatge d'èxit en el batxillerat. Podem afegir que la tendència es confirma si considerem els quartils intermedis: 16 centres tenen un valor afegit positiu per al percentatge d'accés en el tercer quartil, i només 5 en el segon. Una bona part dels "valors afegits" dels instituts és, de fet, el fruit de les desigualtats de nivell escolar dels alumnes en el primer curs, amb una edat i origen donats.

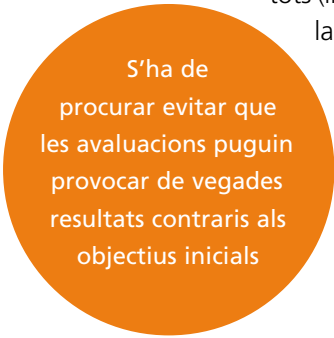
4. Així, doncs, els indicadors de resultats en aquest cas precis penalitzen els instituts més desafavorits en publicar indicadors que permeten pensar que no arriben a assolir els seus objectius, quan en realitat això és per un problema de construcció dels indicadors.
5. Els indicadors de resultats han passat de ser una "eina de pilotatge" a ser simples eines de reproducció de desigualtats entre centres. Cada any es fan públics i reforcen així les "reputacions" dels instituts sobre bases empíriques errònies.

Per tant, una anàlisi del cas francès dels indicadors de resultats mostra que l'avaluació només pot ser una eina de democratització dels sistemes d'ensenyament si són sotmesos a un control metodològic rigorós. En el cas contrari, l'avaluació esdevé una eina de reproducció social de desigualtats entre centres i entre individus.

Els indicadors de resultats penalitzen els instituts més desafavorits perquè permeten pensar que no arriben a assolir els seus objectius

Conclusió

L'objectiu d'aquest text era comprendre el vincle entre avaluació i democratització dels sistemes educatius. És important subratllar el caràcter democràtic de l'avaluació en el sentit que prové de la transparència dels sistemes i s'ha convertit avui en un instrument determinant de les polítiques públiques. M'he esforçat a demostrar que aquest vincle no era mecànic ni estava inscrit en la vocació mateixa de l'avaluació. Es tracta encara d'un procés pròpiament polític que passa pel debat públic i per la capacitat dels ciutadans d'apropiar-se d'aquests resultats de l'avaluació per canviar l'escola. No obstant això, l'exemple francès dels indicadors de resultats ha demostrat també que hem de vigilar tots (investigadors, responsables polítics, ciutadans, etc.) respecte a la construcció i l'ús d'aquestes avaluacions, les conseqüències de les quals poden convertir-se en el contrari d'allò per a què s'havien concebut. I aquesta vigilància s'ha d'estendre a totes les formes d'avaluació, tant a nivell nacional com internacional.



S'ha de procurar evitar que les avaluacions puguin provocar de vegades resultats contraris als objectius inicials

Referències bibliogràfiques

- BAUTIER, E; CRINON, J; RAYOU, P;
ROCHEX, J.-Y. (2007). «Performances en littéracie, modes de faire et univers mobilisés par les élèves: analyses secondaires de l'enquête PISA 2000», *Revue Française de Pédagogie*, n. 157, p. 85-101.
- BERNSTEIN, B. (1975). *Langage et classes sociales*. Paris: Les Éditions de Minuit.
- BOURDIEU, P. (1964). *Les héritiers*. Paris: Les Éditions de Minuit.
- COLEMAN, J.S. et al. (1966). *Equality of Educational Opportunity*. Washington, D.C., U.S.: Government Printing Office.
- ERTL, H. (2006). «Educational standards and the changing discourse on education: the reception and consequences of the PISA study in Germany». *Oxford Review of Education*, vol. 32, n. 5, nov. 2006, p. 619-634.
- FELOUZIS, G. (2005). «Performances et "valeur ajoutée" des lycées: le marché scolaire fait des différences», *Revue Française de Sociologie*, n. 46, 1.
- JENCKS, C. (1978). *L'inégalité*. Paris: Puf.
- KOCH, H. K. (2006). «La reforma educativa alemanya després de l'estudi PISA». *Quaderns d'Avaluació*, n. 5, abril de 2006, p. 55-65.
- LASCOUME, P; LE GALÈS, P. (2004). (dir.), *Gouverner par les instruments*. Paris: Presses de Sciences Po.
- MOREAU, J.; NIDEGGER, C; SOUSSI, A. (2007). «Définition des compétences, choix méthodologiques et retombées sur la politique scolaire suisse», *Revue Française de Pédagogie*, n. 157, p. 43-53.
- NIDEGGER, C. (2005). (coord.), *Pisa 2003: compétences des jeunes romands*. IRDP.
- PASSERON, J.C (1970). *La reproduction*. Paris: Les Éditions de Minuit.
- PISA (2005). *Pisa 2003. Facteurs d'influence sur les résultats cantonaux*. Suisse: Office Fédéral de la Statistique.