

ELS REPTES DE L'AVALUACIÓ DE SISTEMES EDUCATIUS

Josu Sierra Orrantia *Director de l'ISEI-IVEI. Euskadi*

Voldria emmarcar els reptes que planteja l'avaluació de sistemes dins dels cinc reptes dels sistemes educatius europeus que va definir la Comissió Europea l'any 2000 (European Commission, 2000).

Conferència
pronunciada
en el I Seminari
d'Avaluació
Educativa

1. **El coneixement:** l'explosió de la informació demana una redifinició del concepte tradicional de coneixement. Matemàtiques, lectura i ciències són competències clau com a instruments essencials de coneixement i com a base per a l'aprenentatge permanent.
2. **La descentralització:** durant els darrers anys, els sistemes educatius han donat més autonomia i responsabilitat als centres. Fins i tot les avaluacions internacionals (OCDE, Kiel Institute of World Economics) avalen la bondat d'aquesta política.
3. **Els recursos:** l'educació és vista cada vegada més com una inversió. Creix la demanda a l'educació preescolar i augmenten els usuaris de l'educació superior i l'aprenentatge al llarg de la vida.
4. **La inclusió social:** tots els sistemes educatius europeus tenen com a objectiu oferir a la joventut una igualtat d'oportunitats i també preparar-los per a la vida adulta i per sobreviure en una economia globalitzada.
5. **Les dades i la comparabilitat:** el *benchmark*. Són indicadors que, utilitzats de manera diagnòstica i formativa, ens permeten comparar-nos per informar la política educativa i conèixer les tendències. El repte de la comparació és crear un clima obert i positiu de diàleg tenint els indicadors com a punt de partida.

Voldria destacar, sobretot, el cinquè repte, ja que aquí hi ha la clau, avui dia, per a la millora del nostre sistema educatiu.

El punt de partida: un sistema centralitzat que desconex bàsicament els seus resultats

Fins ara, el nostre sistema educatiu ha estat un sistema preocupat per l'*input* (el currículum) normatiu i centralitzat. Tots els centres han de seguir els decrets curriculars, i la gestió dels recursos correspon a l'Administració. No obstant això, es desconexien els resultats del sistema i si tots els centres seguien o no els currículums, en quina mesura

ho feien, etc. Les decisions de política educativa es prenen o bé intuïtivament o bé a partir d'opinions.

L'any 2001, quan es va crear l'Institut Basc d'Avaluació i Investigació Educativa (ISEI-IVEI), només teníem dades de resultats del nostre sistema a partir de la nostra participació en avaluacions estatals coordinades per l'INECSE (Institut Nacional d'Avaluació i Qualitat del Sistema Educatiu):

- Avaluació de l'educació primària 1995
- Avaluació del diagnòstic del sistema educatiu 1997

L'aportació de les avaluacions estatals

Aquesta aportació era absolutament insuficient per conèixer els nostres resultats i construir indicadors. Només ens proporcionen la mitjana estatal, la pròpia i la tendència en anys posteriors. A més, sols mesuren el currículum comú, cosa que obliga a fer proves complementàries.

En realitat, després de la filtració dels resultats de l'avaluació de primària 1995 per part de comunitats autònomes es va produir un ús tendencios de l'avaluació, que va ser emprada per menysprear determinades comunitats des de la premsa. Després d'aquest incident, l'INECSE va decidir no fer mostres representatives de les comunitats en els seus estudis.

Les avaluacions estatals aporten dades insuficients per conèixer els resultats del sistema educatiu basc

Els informes publicats per l'INECSE només parlen de resultats de l'Estat i no estan orientats a la política educativa, que és diferent en cada comunitat. Avui dia, es dona la paradoxa que moltes comunitats no mesuren els seus resultats amb aquestes proves i que el MEC sí que ho fa, en un àmbit estatal. Quin sentit té això si la finalitat de l'avaluació és la millora?

Les proves estatals només avaluen el currículum comú, basat en un acord entre les comunitats que van aportar els ítems i l'INECSE.

Tot això ens va acabar de convèncer que les nostres avaluacions han de basar-se en proves internacionals.

La nostra aposta per l'avaluació internacional com a font d'indicadors i d'informes sobre el sistema, orientats a la política educativa

A partir de l'estudi PISA (Programa Internacional per a l'Avaluació de l'Alumnat) 2000, l'avaluació internacional s'ha convertit en un dels principals motors de la millora dels sistemes educatius i el Simposi PISA de Berlín de l'any 2002 en va ser la prova més palpable. Vet aquí alguns dels seus postulats:

- Impulsar a la vegada bons resultats i equitat no és un somni. És possible i així ho fan alguns països.
 - Un paper essencial del sistema educatiu és el de compensar les diferències socials d'origen de l'alumnat i donar igualtat d'oportunitats a tothom.
 - Existeix un debat enriquidor al voltant de PISA amb l'objectiu de millorar.
 - Sabem que el nivell socioeconòmic de la família és molt important a l'hora d'explicar el rendiment de l'alumnat. Però no és determinant. La seva importància varia d'uns països a uns altres.
 - Alguns sistemes educatius compensen les diferències socioeconòmiques, mentre que altres les multipliquen.
- També són interessants algunes de les seves conclusions, publicades fa poc per l'OCDE a *Messages from PISA 2000* (OCDE, 2004, p. 19):

- Els sistemes educatius que tenen èxit han augmentat l'autonomia dels centres escolars. A més, l'han combinada amb la devolució de la responsabilitat als centres.
- Els sistemes que tenen èxit estan compromesos amb la monitorització del rendiment de l'alumnat i del sistema. Encara que no existeix una metodologia única, els millors països monitoritzen el rendiment dels seus estudiants i dels seus centres.
- El mètode usat per un sistema educatiu per ajudar l'alumnat amb baix rendiment és crucial per a la millora dels resultats. A més, les diferències entre centres tendeixen a augmentar en els sistemes no comprensius que tenen una segregació precoç d'alumnat en itineraris diferents.

L'avaluació internacional és un dels principals motors de millora dels sistemes educatius: per això apostem pels estudis PISA

Encara que ja havíem decidit participar en el PISA 2003, la nostra assistència al Simposi PISA de Berlín de l'any 2002 ens va acabar de convèncer de la bondat de l'aposta per l'avaluació internacional. Per això vam decidir participar amb més convenciment en els estudis PISA i TIMSS (Estudi Internacional de Tendències en Matemàtiques i Ciències) de l'any 2003.

Necessitàvem indicadors europeus i mundials, comparar-nos amb aquells països amb els quals compartim experiències en comú i competim en els mercats. Vam trobar que les avaluacions de l'OCDE (Organització per a la Cooperació i el Desenvolupament Econòmic) i la IEA (Associació Internacional per a l'Avaluació del Rendiment Educatiu) ens oferien:

- Tecnologia d'avaluació del més alt nivell
- Proves acceptades mundialment
- Comparabilitat internacional
- Dades del nostre sistema per a la política educativa
- Indicadors europeus en competències clau

Pisa 2003: un gran avenç

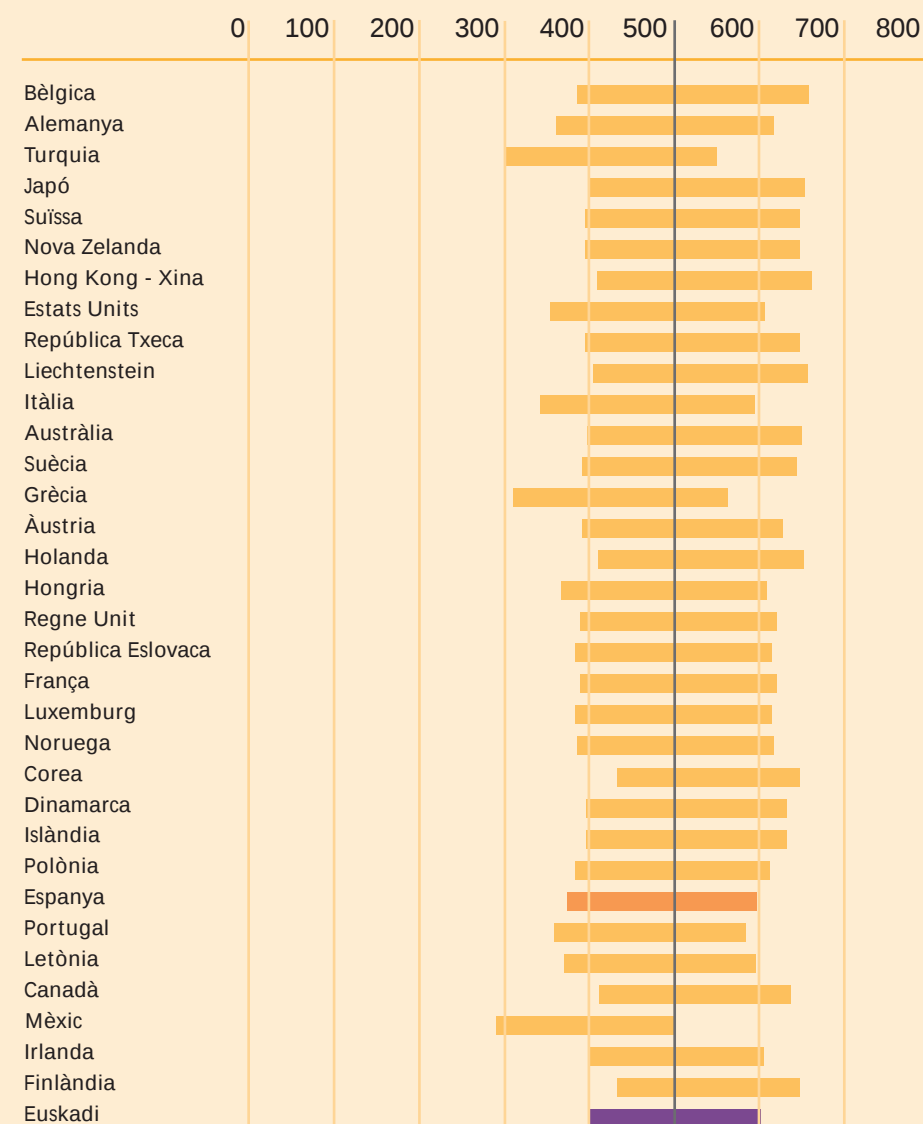
El PISA 2003 ens permet "situar-nos", conèixer els nostres resultats en competències clau (indicador europeu), veure els nostres punts febles i les característiques del nostre sistema.

Si fem una síntesi dels resultats obtinguts per l'alumnat d'Euskadi, veurem que:

- En tres de les quatre competències clau ens situem en la mitjana de l'OCDE: coneixement matemàtic, lectura i resolució de problemes.
- En coneixement científic estem per sota de la mitjana dels països de l'OCDE. Aquesta ha de ser una prioritat en els propers anys: millorar en coneixement científic.
- El nostre sistema és d'una gran equitat: hi ha molt poca dispersió en el rendiment de l'alumnat. El nostre percentatge d'alumnat de baix rendiment és tan sols del 16%, enfront del 23% d'Espanya, per exemple.
- Tenim menys alumnat de rendiment baix que l'OCDE, però ens falta excel·lència, és a dir, que tenim molt poc alumnat excel·lent en matemàtiques.

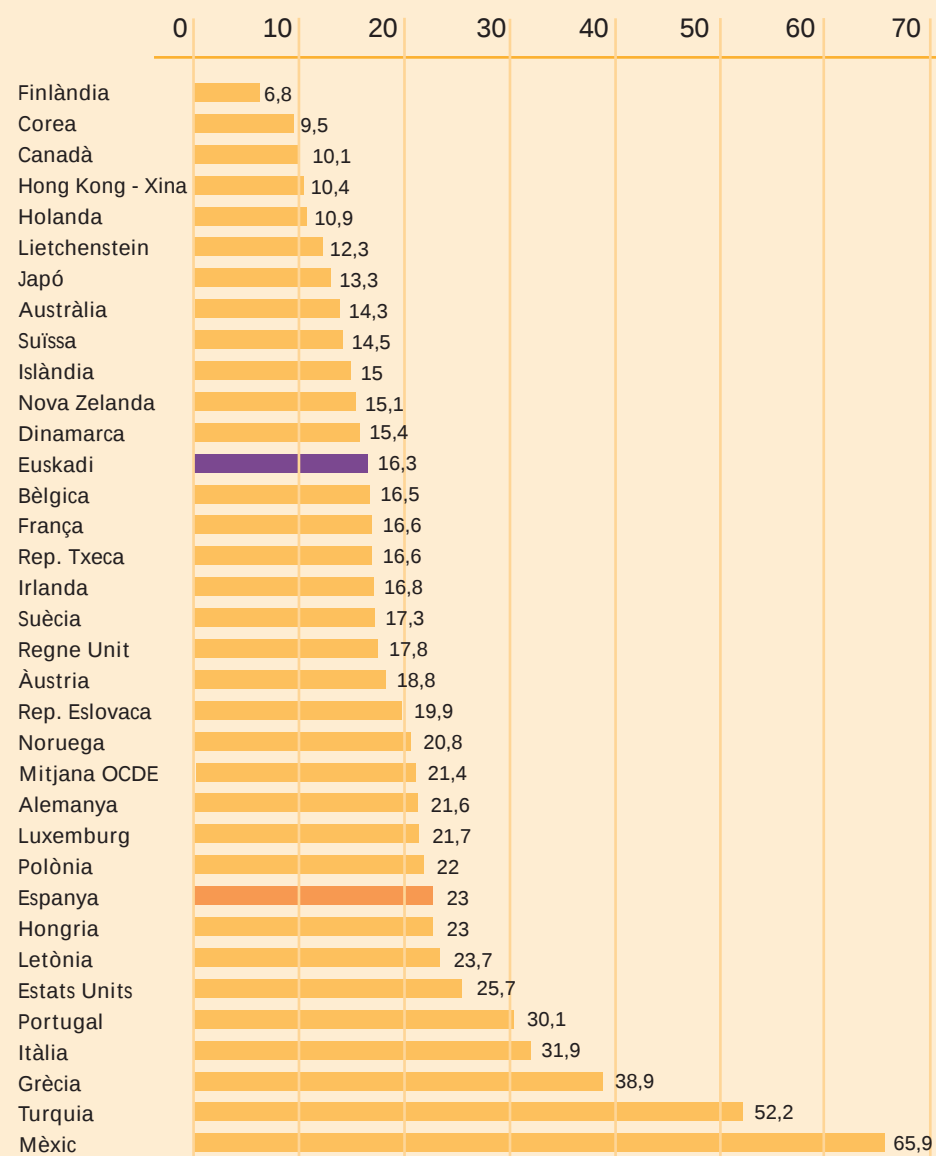
El gràfic següent mostra que **la nostra equitat relativa és la més gran de l'OCDE i que l'absoluta és també important:**

Gràfic 1 ■ Diferència de puntuació entre els alumnes millors i pitjors (equitat). Matemàtiques



Euskadi té la diferència més petita entre alumnat bo i dolent de tota l'OCDE:

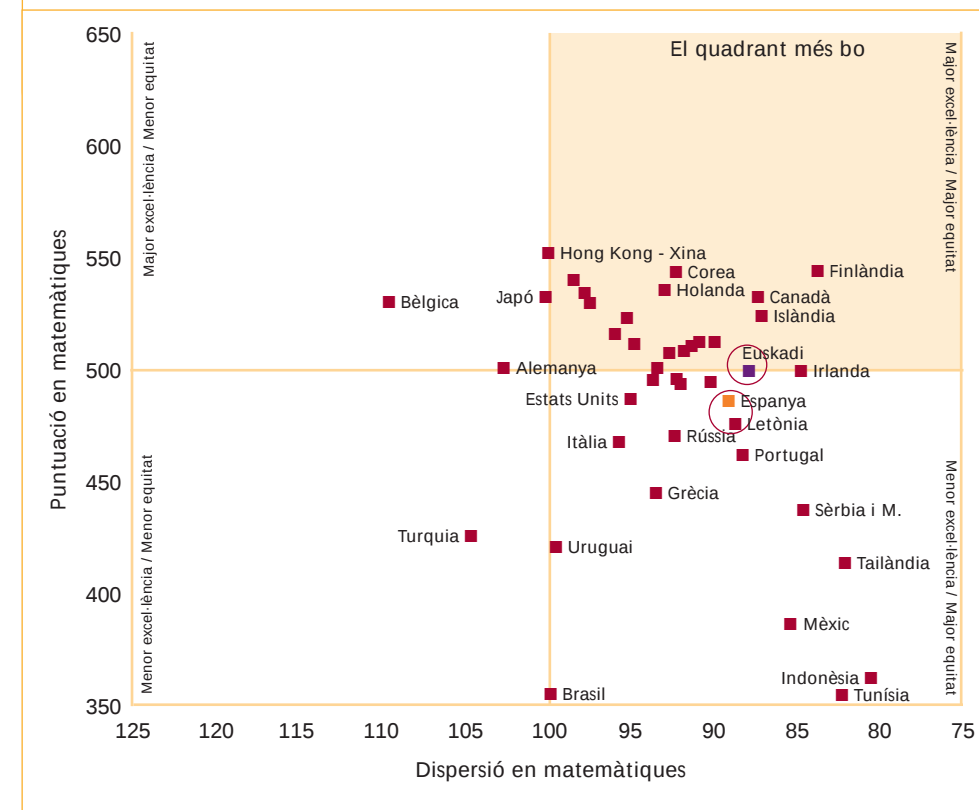
Gràfic 2 ■ Equitat absoluta: % d'alumnes en els nivells 0 i 1. Matemàtiques



A més a més, el percentatge d'alumnat en els nivells 1 i menys 1 de PISA és del 16,3%, el lloc 13 de l'OCDE.

El gràfic següent mostra la **relació entre rendiment i equitat** (representada per la desviació típica):

Gràfic 3 ■ Equitat Pisa 2003



Si posem els dos paràmetres en relació, veurem que Euskadi se situa en una bona posició, que encara és millorable.

La **relació entre el rendiment i el PIB per càpita** es posa de manifest en el gràfic següent, realitzat per l'INECSE. Si hi comparem el nostre nivell de riquesa amb la puntuació que hem obtingut en el PISA 2003, veurem que ens situem per sobre de la recta de regressió, és a dir, del nostre rendiment esperable.

Gràfic 4 ■ Relació entre el rendiment i el PIB per càpita



Altres resultats importants sobre el sistema educatiu

A Euskadi, la variable “nivell socioeconòmic i cultural de l'alumnat” i la variable de la “mitjana del nivell socioeconòmic i cultural del centre” tenen una baixa influència en els resultats de l'alumnat. El seu nivell d'importància és equivalent al que tenen a Finlàndia o al Canadà, països model per a nosaltres.

Però a més a més, en matemàtiques tenim una de les diferències entre centres més baixes de tots els països que han participat en l'avaluació. Darrere d'Islàndia, Finlàndia, Noruega i Suècia, la nostra diferència de rendiment entre centres és la més baixa de l'estudi PISA.

Timss 2003. Avaluació de les matemàtiques i les ciències a segon d'ESO

Euskadi també ha participat en aquesta avaluació impulsada per l'IEA i aquest cop ho ha fet en solitari, ja que Espanya no hi ha participat.

En resum, els nostres resultats en matemàtiques mostren la mateixa puntuació que Nova Zelanda, Eslovènia i Itàlia i estem per sobre de la mitjana internacional. En ciències tenim la mateixa puntuació que Noruega, Itàlia, Israel i Bulgària, també per sobre de la mitjana internacional. Es reafirmen, doncs, els resultats de l'estudi PISA pel que fa a l'equitat del sistema educatiu basc.

Els resultats en matemàtiques i ciències de l'estudi TIMSS 2003 confirmen els resultats del PISA sobre l'equitat del sistema educatiu basc

En definitiva, en el nostre sistema educatiu hi ha un abans i un després a partir de 2004. Gràcies a la nostra aposta per l'avaluació internacional estem en condicions de prendre decisions sobre propostes per a la millora del rendiment del nostre alumnat i sabem moltes més coses que abans sobre el nostre sistema.

Límits de l'avaluació: mostres, censos, proves diagnòstiques

Un dels problemes de l'avaluació de sistemes educatius és la limitació dels estudis mostrals. Els estudis basats en una mostra de centres i d'alumnes estan limitats perquè només serveixen per donar una visió general del sistema. Permeten detectar problemes generals de rendiment, però moltes vegades no n'expliquen les causes. Les variables explicatives es queden en un pla general de nivell socioeconòmic i cultural, de formació dels pares i mares, de l'estructura del sistema educatiu, etc. Una de les limitacions més grans dels estudis mostrals és que no permeten parlar de centres, i encara menys d'alumnes de determinats centres.

Davant de la publicació dels resultats, els centres escolars no s'hi consideren involucrats i no es pot fer un “mapa” del rendiment del sistema o una “base de dades” que permeti actuar allà on realment es necessiti en termes de grups, de centres o d'alumnat.

Aquesta limitació no és inherent a les avaluacions internacionals, sinó que ho és a qualsevol avaluació. Les avaluacions internacionals permeten ser aplicades a tot l'alumnat d'una edat o d'un curs, encara que això, naturalment, encareix l'anàlisi de dades.

La **tendència actual en altres països** és clarament favorable a l'obligació de retre comptes (*accountability*). Es combinen proves úniques, gairebé sempre basades en estàndards, que s'apliquen a tot l'alumnat amb una devolució de l'autonomia i la responsabilitat als centres educatius.

Aquests sistemes utilitzen proves censals, diagnòstiques o no, que:

- Permeten conèixer on són els problemes i quins són. Moltes vegades les avaluacions mostrals detecten un problema però no en poden explicar les causes.
- Tanmateix, l'efecte de les proves diagnòstiques és que es converteixen “de facto” en el currículum i l'avaluació passa a convertir-se en norma per al professorat. Per evitar això, és important que:
 - Els estàndards en què es basen les proves no siguin reduccionistes ni simplificadors.
 - Les proves no abusin dels ítems de resposta múltiple i continguin preguntes obertes, de construcció complexa, d'interpretació, etc.
 - Les proves siguin el fruit d'un gran consens.

Les proves censals potser permetrien calcular el valor afegit, si s'apliquessin durant diversos anys seguits a la mateixa cohort i ja existís una prova estandarditzada de referència. Aquesta és la forma més justa d'avaluar els rendiments: no el rendiment que l'alumnat aconsegueix en les proves (rànquings), sinó la millora que s'ha produït d'un any a un altre.

Les proves diagnòstiques censals

Sembla que les proves diagnòstiques de 4t de primària i 2n de secundària proposades en el document “Propuesta para el debate”, del Ministeri d'Educació i Ciència, serien similars, en alguna mesura, a les que acabem d'esmentar. La LOCE ja propugnava la realització de proves diagnòstiques, però no hi havia prevista una devolució d'autonomia i responsabilitat als centres. Al meu parer, el control sense l'autonomia no funcionarà.

Hi ha una altra qüestió clau. Si “l'avaluació sense conseqüències no és avaluació”, com diuen a Alberta [província del Canadà capdavantera en investigació avaluativa], és evident que cal preveure mecanismes de millora una vegada coneguts els resultats, i això

val per a totes les avaluacions. No es pot deixar tot com estava després de l'avaluació, sobretot en aquells centres que obtinguin uns resultats més baixos.

El problema és que l'estructura actual del sistema educatiu no permet més que de manera limitada la intervenció asimètrica en alguns centres i en d'altres no, la qual cosa és especialment greu en qüestions de gestió de personal. Ens hauríem de preguntar si el sistema funcional actual, tal com està establert en la xarxa pública, és compatible amb aquest tipus de rendició de comptes. En aquest sentit, a França disposen d'un sistema més descentralitzat tenint en compte el tipus de direcció de centres de què disposen i les anomenades zones educatives prioritàries (ZEP), amb una assignació de recursos molt especial.

És clar que s'ha de tenir en compte, també, l'ús diferent de les proves censals que es fa en països distints. Mentre que al Canadà o a Suècia, per posar un exemple de països que tenen un sistema absolutament descentralitzat del ministeri, els resultats es fan públics per centre i la ciutadania els pot consultar per Internet, amb les conseqüències que això comporta en la tria de centre per part dels pares i mares, en països com França, més centralitzats, el Ministeri d'Educació es queda amb una mostra i fa públic un informe anual.

Alemanya ha estat l'últim país que ha elaborat un sistema de rendició de comptes com a efecte de l'estudi PISA 2000 i ha començat l'aplicació de proves censals el setembre de 2004. Al Canadà o a Suècia, el ministeri disposa cada any d'una “base de dades” amb tots els resultats de tots els centres i hi actua en conseqüència allà on cal. A França es pretén que el mateix professorat i el centre treguin les seves pròpies conclusions amb les referències que se'ls facilita sobre l'avaluació (vegeu <http://educ-eval.education.fr/eval-diag.htm> i <http://www.evace26.education.gouv.fr/>).

Hi ha efectes col·laterals previsibles, però. Els centres que obtinguin bons resultats en les proves diagnòstiques els faran públics amb la finalitat de demostrar als pares i mares el seu alt nivell de rendiment, la qual cosa acostuma a exigir fer públics alguns informes que expliquin les diferències entre centres arran del nivell socioeconòmic de l'alumnat o del nivell cultural de les famílies.

Una altra consideració a fer és que aquestes proves han de tenir estàndards adaptats a la realitat de cada comunitat autònoma, encara que es pot establir un mecanisme de comparació o d'equivalència. Un altre problema és el de la recollida de les dades

L'avaluació sense conseqüències no és avaluació.

Les proves d'avaluació censals permeten detectar quins són els problemes i on són amb més exactitud que no pas les mostrals.

(normalment a través d'Internet), el de l'anàlisi d'una quantitat ingent de dades en un temps raonable i el del finançament de tot plegat. Per dedicar-hi els recursos necessaris cal creure en l'avaluació i fer una aposta decidida per l'avaluació com a pas previ a la millora dels sistemes educatius.

Un dels problemes de les avaluacions diagnòstiques per a 4t de primària i 2n d'ESO proposades en el debat educatiu és que es proposa continuar amb les avaluacions estatals i les internacionals al mateix temps.

Si ens fixem en la taula següent ens adonarem que això implica realitzar avaluacions redundants, amb el perill de dilapidar els recursos i de trobar-nos amb resultats dispars que serien difícils d'explicar.

Avaluacions proposades	Nivell	Què mesuren	Nivell tecnològic	Valor afegit
PISA 2006	15 anys (8è, 9è i 10è)	Competències clau: ciències, matemàtiques, lectura	Mostreig i anàlisi de dades al més alt nivell	Proves úniques. Comparabilitat internacional i OCDE. Informe orientat a la política educativa
TIMSS 2007	4t, 8è	Matemàtiques, ciències	Mostreig i anàlisi de dades al més alt nivell	Proves úniques. Comparabilitat internacional
PIRLS 2006	4t	Lectura	Mostreig i anàlisi de dades al més alt nivell	Proves úniques. Comparabilitat internacional
Proves diagnòstiques	4t, 8è	—	—	—
Avaluació estatal de primària	6è	Currículum comú per acord entre comunitats: matemàtiques, llengua, coneixement del medi, anglès	Pobre	Comparabilitat amb la mitjana de l'estat. Tendència
Avaluació estatal de secundària	10è	Currículum comú per acord entre comunitats: matemàtiques, llengua, coneixement del medi, anglès	Pobre	Comparabilitat amb la mitjana de l'estat. Tendència

Que les futures proves diagnòstiques coincideixin amb les avaluacions internacionals en el mateix nivell pot esdevenir una duplictat innecessària i ho serà molt més encara si les avaluacions estatals coincideixen amb avaluacions internacionals al final d'algunes etapes. Seria més pràctic sotmetre's al dictamen de PISA pel que fa a competències clau, tot ampliant l'avaluació amb altres proves en el cas que es consideri necessari.

En definitiva, sembla que caldria reflexionar sobre què volem aconseguir realment amb l'avaluació i què podem dur a terme.

Per què no aprofitem les avaluacions internacionals per avaluar els finals d'etapa (fins i tot com a proves diagnòstiques)?

En resum i per acabar, m'agradaria remarcar les dues idees que he defensat al llarg d'aquestes pàgines. En primer lloc, que les avaluacions diagnòstiques censals tenen una potència avaluativa molt gran però que, al meu parer, s'han de combinar amb l'autonomia i responsabilitat dels centres sobre els resultats obtinguts. En segon lloc, que les avaluacions internacionals són la nostra aposta actual, fins i tot com a avaluació de final d'etapa. El PISA, per exemple, té una opció anomenada de 10è nivell, que permet fer una mostra de 4t d'ESO i convertir-la en avaluació de la secundària obligatòria.

Referències bibliogràfiques

European Commission. Directorate-General for Education and Culture (2000) "European Report on the quality of School Education". <http://europa.eu.int/comm/education/policies/educ/indic/rapinen.pdf>

OCDE (2004) "Messages from PISA 2000". París <http://www.pisa.oecd.org/dataoecd/31/19/34107978.pdf>