

# L'avaluació del sistema educatiu

## Notes sobre el diagnòstic del sistema educatiu 1997 de l'INCE

Màrius Martínez Muñoz<sup>1</sup>

Universitat Autònoma de Barcelona. Departament de Pedagogia Aplicada  
08193 Bellaterra (Barcelona). Spain

---

### Resum

Durant el curs 1996-1997 l'Instituto Nacional de Calidad y Evaluación (INCE) va desenvolupar una avaluació del sistema educatiu en els nivells dels alumnes que tenien 14 i 16 anys. Aquest estudi va ser àmpliament difós pels mitjans de comunicació, que es feien ressò de l'impacte que va tenir en cercles educatius.

Aquest article repassa els antecedents de l'estudi des d'una panoràmica tant nacional com internacional; també en descriu, a grans trets, les característiques i alguns dels resultats més significatius als quals van arribar les cinc comissions d'estudi de l'INCE: resultats acadèmics, plans d'estudi, funcionaments dels centres, perfil professional i relacions amb la societat. La descripció del disseny de l'estudi així com del desenvolupament del treball i de les conclusions es combina amb comentaris i reflexions crítiques que els matisen i els valoren.

**Paraules clau:** avaluació del sistema, metodologia d'avaluació, disseny d'avaluació, resultats acadèmics, mostres.

---

### Resumen

Durante el curso 1996-1997 el Instituto Nacional de Calidad y Evaluación (INCE) llevó a cabo una evaluación del sistema educativo en los niveles de edad de 14-16 años. Este estudio fue ampliamente difundido por los medios de comunicación haciéndose eco del impacto que tuvo en los núcleos educativos.

Este artículo repasa los antecedentes del estudio desde una panorámica tanto nacional como internacional. También describe, a grandes rasgos, las características del mismo y alguno de los resultados más significativos a los que llegaron las cinco comisiones de estudio del INCE: resultados académicos, planes de estudio, funcionamiento de los centros, perfil profesional y relaciones con la sociedad. La descripción del diseño del estudio así

1. L'autor, professor de l'àrea de MIDE del Departament de Pedagogia Aplicada a la Facultat de Ciències de l'Educació de la UAB, ha participat en el *Diagnòstic del Sistema Educativo* promogut per l'INCE durant el curs 1996-1997, com a membre de la comissió especialitzada V que va treballar les relacions escola i societat. A proposta del Consell Superior d'Avaluació del Sistema Educatiu del Departament d'Ensenyament de la Generalitat de Catalunya amb el qual ha col·laborat en l'avaluació del clima escolar a centres de Secundària (com a codirector) i en un projecte internacional d'avaluació dels programes Comènius I a centres d'Educació Primària i Secundària.

como el desarrollo del trabajo y las conclusiones se combinan con comentarios y reflexiones críticas que se matizan y valoran.

**Palabras clave:** evaluación del sistema, metodología de evaluación, diseño de evaluación, resultados académicos, muestras.

### Abstract

---

During the academic course 1996-1997 the National Institute for Quality and Evaluation (INCE) has developed an evaluation of the educational system studying the levels in which the pupils were 14 and 16 years old. This study has been largely diffused through the mass media reflecting the big impact that it has had in educational environments.

This article overviews the antecedents of the study both, from a national and an international point of view. It describes also its general characteristics and some of the most relevant results that the five working commissions had reached concerning: educational outcomes, professional profile, curriculum, the functioning of the institutions and their relationships with society. The description of the aforementioned topics is combined with some critic reflections and comments.

---

### Sumari

#### 1. Introducció: el context de l'avaluació del sistema

L'avaluació del sistema educatiu espanyol, desenvolupada durant el curs 1996-1997, ha generat un notable rebombori en cercles educatius reflectit en els mitjans de comunicació. Una avaluació d'aquest abast en la qual van ser enquestats més de 40.000 alumnes, 11.000 famílies i 6.000 docents, en relació a diversos aspectes de l'educació als 14 i als 16 anys a l'Estat espanyol, ha merescut l'interès no solament de professionals de l'educació sinó també dels polítics i de l'opinió pública en general. El fet que els seus resultats tinguin un gran impacte en el conjunt de la societat porta a establir-ne mecanismes d'avaluació, de seguiment i sobretot de millora.

Aquest estudi avaluatiu s'hauria d'entendre i situar tant en un context de canvi i de millora de la qualitat de l'educació com també des d'una òptica internacional de desenvolupament de sistemes d'avaluació, seguiment i millora dels sistemes educatius, que són accions complementàries a les avaluacions internes i externes dels centres educatius i que haurien de tendir, totes elles, a la creació d'una cultura avaluativa de millora de l'educació amb tots els seus protagonistes.

### *El context de l'avaluació*

L'avaluació ha esdevingut, en els darrers anys, un tòpic recurrent en la literatura educativa i cada cop més en la pràctica de professionals i institucions dels diferents nivells educatius. Els textos sobre avaluació i les accions avaluatives tenen cada dia més presència i importància en la realitat educativa, tant a Catalunya i Espanya com també en el context internacional.

### *L'avaluació de l'educació en el context internacional*

En el panorama internacional i en un sentit ampli, l'avaluació de l'educació ha estat també una pràctica que compta amb una certa acumulació d'experiències (es pot recordar per exemple l'informe mundial signat per P. H. Coombs del Consell Internacional per al Desenvolupament de l'Educació (ICED) sobre la crisi mundial de l'educació de la dècada dels seixanta o el darrer informe de J. Delors sobre l'educació al segle XXI, encarregat per la UNESCO. Aquests informes si bé apleguen una notable quantitat d'informació i dades tant quantitatives com qualitatives, han de tenir per força, i donat l'abast de les seves reflexions, conclusions i propostes, un caràcter general que serveix per inspirar polítiques regionals i estatals i que dibuixa en grans termes les tendències que en matèria d'educació semblen detectar-se.

Hi ha, però, altres iniciatives internacionals que mereixen també esment per la seva importància i abast. Així la *International Association for the Evaluation of Educational Achievement* (IEA) ha desenvolupat des del 1964 estudis sobre el rendiment en matemàtiques i en ciències la darrera edició dels quals es va desenvolupar el 1994 donant lloc al tercer estudi internacional en matemàtiques i ciències<sup>2</sup> en el qual van participar 45 països d'arreu del món i més de 500.000 alumnes de 9 i 13 anys.

Cal destacar també entre aquestes iniciatives internacionals les impulsades per l'OCDE (Organització per a la Cooperació i el Desenvolupament Econòmic), que publica periòdicament informes i estadístiques sobre educació en aquesta perspectiva internacional. Concretament mereixen especial esment el projecte INES (*International Indicators of Education Systems*), que té com a objectiu la creació d'indicadors educatius sobre els sistemes educatius dels seus països membres que inclouen indicadors comparatius internacionals del rendiment escolar dels alumnes. Les dades proporcionades per aquests indicadors són objecte de difusió en l'informe que amb el títol *Education at a glance* publica anualment l'OCDE.

L'Estat espanyol i les comunitats amb competències en educació han participat en aquests estudis internacionals dels quals se n'han anat publicant els resultats obtinguts.

2. Aquest estudi és conegut per l'acrònim TIMSS (Third International Mathematics and Science Study).

## *L'avaluació del sistema a Espanya i a Catalunya*

A l'Estat espanyol s'han desenvolupat nombrosos estudis avaluatius del sistema educatiu, especialment a partir de la dècada dels anys 70 amb el concurs de diferents organismes que els han anat signant. Es poden destacar els llibres blancs per a la reforma del sistema que es van publicar el 1969 i el 1989 impulsats per les autoritats educatives estatals. Es poden destacar també els informes sobre l'Educació Secundària (1981), l'informe sobre la política educativa espanyola (aquest desenvolupat per l'OCDE el 1986) o sobre els centres educatius i la qualitat de l'ensenyament (1994) que es van anar desenvolupant. Aquell mateix any l'INCE (creat aleshores) va fer un estudi sobre el sistema educatiu, en aquell cas de l'Educació Primària, amb una mostra de 10.000 alumnes.

L'INCE destaca que els estudis que versen sobre diversos aspectes del sistema sumen més de 300: l'estructura general del sistema educatiu, la gestió del sistema, el currículum, el professorat, les desigualtats en educació, l'avaluació i la inspecció, el rendiment dels alumnes, entre altres (Ministerio de Educación y Cultura: 1998b).

A Catalunya es pot destacar l'estudi publicat el 1991 que, amb el títol d'*Avaluar per innovar*, va servir per avaluar els plans d'experimentació de la reforma educativa en cinc àmbits: curricular, d'orientació, d'organització i gestió, de formació de l'ensenyant i de costos econòmics. També es poden destacar els darrers informes d'avaluació del sistema publicats pel Consell Superior d'Avaluació del Sistema Educatiu: sobre indicadors del sistema educatiu a Catalunya (1997), sobre el clima escolar als centres de Secundària de Catalunya (1998) i, en fase d'elaboració, un conjunt d'informes sobre el diagnòstic del sistema educatiu però utilitzant només les dades de Catalunya.

Tot aquest esforç avaluatiu es genera arran de la legislació que en matèria d'educació s'ha anat desenvolupant, especialment a partir de la LOGSE que ja en el seu títol IV esmenta l'avaluació del sistema educatiu com un dels factors que afavoreixen la qualitat i la millora del sistema. L'estudi publicat per l'INCE vol ser un diagnòstic del sistema educatiu.

## **2. El disseny de l'estudi**

### *Característiques generals*

L'estudi objecte de comentari en aquest article va ser un afegit, una ampliació, del pla d'actuació de l'INCE per al període 1994-1997. I va respondre a l'interès de la institució de fer una avaluació de la situació del sistema educatiu en la franja 14-16 anys (corresponent al següent tram d'educació obligatòria després de l'ensenyament primari que fou objecte de l'anterior avaluació per part d'aquest mateix organisme). El motiu d'ampliar el pla de quatre anys va ser fer-lo abans del traspàs de competències en matèria d'educació a totes les autonomies que encara no les tenien transferides, buscant a més que fos representatiu de tot l'Estat espanyol. No es va prendre com a criteri el nivell educatiu perquè

**Quadre 1.** Temes tractats per les 5 comissions de l'avaluació.

Comissió I:	Rendiment acadèmic de l'alumnat de 14 i de 16 anys
Comissió II:	Plans i mètodes d'ensenyament i aprenentatge
Comissió III:	Funcionament dels centres escolars
Comissió IV:	La funció docent
Comissió V:	Escola i societat

això hauria deixat fora de l'estudi molts alumnes en funció del sistema que estiguessin cursant<sup>3</sup>.

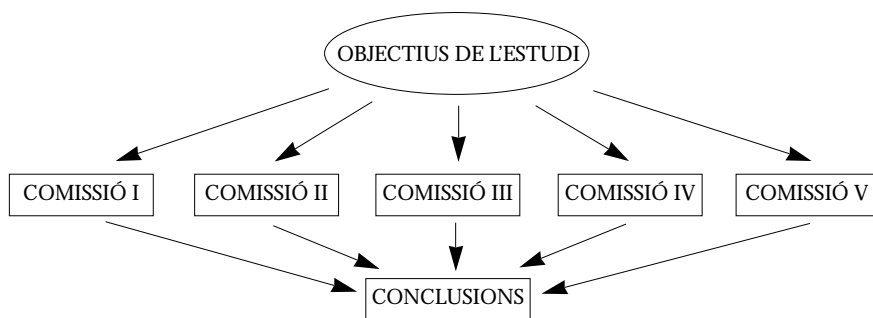
Aquest estudi va ser dissenyat pensant a avaluar no només els aspectes tradicionals en aquestes accions que són els resultats acadèmics<sup>4</sup>, sinó buscant altres elements igualment importants per al sistema. D'aquesta manera es van dissenyar cinc focus d'estudi que serien encarregats a cinc comissions de treball diferents (quadre 1).

Cada comissió estava integrada per un president, designat per l'INCE, i nou representants de les comunitats amb plenes competències i altres proposats pel MEC representant el territori encara sota les seves competències (l'anomenat *Territori MEC*)<sup>5</sup>.

El disseny en cinc comissions amb encàrrecs genèrics és un punt clau de l'estudi perquè determina el curs del procés i en bona mesura les possibilitats dels resultats finals, com comentarem més endavant.

D'altra banda, cal fer esment també de la dificultat d'avaluar un sistema educatiu descentralitzat. Es pot pensar si l'avaluació respon a una avaluació d'un sistema o bé si es tracta d'avaluar un sistema descentralitzat que no és exactament el mateix. Cal considerar el fet de tenir un sistema educatiu que està organitzat en tres nivells de concreció i que està gestionat i regulat en darrer terme per administracions diferents amb diferents programes polítics que necessàriament afecten l'articulació del sistema en cada autonomia. Aquest fet complica la determinació de quin és el currículum comú (que per molt que està explícit al primer nivell de concreció dóna lloc a concrecions diferents).

3. En el moment de desenvolupar l'estudi estaven coexistint l'antiga Formació Professional, el BUP i l'ESO. Aquesta barreja de línies curriculars és un dels motius que justifiquen no poder parlar de l'avaluació de l'ESO. Tot i això el títol genèric que figura a tots els informes d'aquest estudi és «Diagnòstic del Sistema Educativo. La escuela secundaria obligatoria» que no s'hauria de confondre amb l'Ensenyament Secundari Obligatori.
4. Les àrees avaluades van ser llengua i literatura (catalana i castellana a Catalunya), matemàtiques, i al territori MEC ciències socials i naturals. Aquestes dues últimes no s'han tingut en compte en aquest article.
5. Totes les comunitats van enviar representants a totes les comissions tot i que la Junta d'Andalusia, al·legant que estava desenvolupant un diagnòstic propi que feia inviable la seva participació (MEC: 1998b: 11), no va voler participar en l'estudi i només va presentar-se en alguna comissió d'observadors. La comunitat de Canàries va desmarcar-se, també, d'algunes comissions, i va participar finalment en les comissions IV i V.



**Figura 1.** Esquematització del disseny d'avaluació utilitzat.

Aquesta discussió encara pot anar més enllà (com de fet va succeir en les comissions) quan fins i tot dins d'una mateixa autonomia poden haver interpretacions diferents sobre la importància o pertinença de considerar uns aspectes o uns altres. Encara que, com s'explicarà més endavant, els models d'interpretació de resultats permeten mitigar la qüestió, aquesta no queda del tot resolta.

En certa manera, aquesta organització porta a un disseny de l'avaluació fragmentat, lineal i additiu. Es podria representar a la figura 1.

La solució a aquesta qüestió no és fàcil, però cal tenir-ho en compte en el moment de dissenyar l'estudi en el seu conjunt. Això afectarà consegüentment els possibles grups de treball, les mostres, els instruments i el tipus de tractament de la informació i, el que és més important, les conclusions i les implicacions per a la pràctica de la millora i optimització del sistema. L'informe d'avaluació integra parcialment els resultats en les conclusions finals en forma de 109 conclusions i recomanacions que relacionen, en part, les conclusions parcials de cadascuna de les comissions especialitzades.

### *La perspectiva de l'estudi*

Aquest estudi s'emmarca, fonamentalment, en una perspectiva empírico-analítica en què es tracten i s'interpreten dades quantitatives encara que després es facin interpretacions qualitatives i que, en alguns casos, també s'inclouin algunes dades qualitatives (obtingudes en les taules rodones i en els grups de diagnòstic, tal com s'explicarà més endavant).

Es consideren fonamentalment els resultats acadèmics i les valoracions de diferents aspectes de l'educació, en una perspectiva d'objectivitat i científicitat. La concepció és positivista, de recerca de l'objectivitat, d'utilització de mecanismes per a l'estratificació i aleatorització en la selecció de les mostres. Els instruments es construeixen amb criteris de validesa i fiabilitat. Els dissenys són estructurats i l'explicació de resultats es fa a partir de procediments estadístics.

Es posa l'èmfasi en els resultats per sobre d'altres consideracions. El gruix dels estudis esmentats a la introducció adopten criteris similars que es justifi-

quen per l'abast dels processos avaluatius endegats. Quan es tracta amb regions i estats la realitat, que és molt més rica i diversa, es redueix per poder ser tractada tot assumint que aquesta reducció fa perdre molts matisos i relacions que difícilment poden quantificar-se i operativitzar-se.

Els avaluadors són elements externs a la realitat avaluada i les conclusions, per tant, se situen en un pla tècnic que cal concretar en accions polítiques per part de les administracions competents.

Les audiències reals dels informes són, fonamentalment, les administracions, els professionals de l'ensenyament, les famílies i els mateixos alumnes. Tots ells, a diferents nivells, tenen capacitat real d'incidir sobre els aspectes avaluats que tinguin menys satisfactorietat.

### 3. La metodologia de l'estudi

#### *El mostreig*

El mostreig és un aspecte rellevant quan es vol fer una avaluació d'una població massa extensa com per poder-la abastar en la seva totalitat.

El quadre número 2 recull els principals elements il·lustratius dels diferents processos de mostreig que es van desenvolupar en l'informe<sup>6</sup>.

Com es pot apreciar al quadre 2, el procés de mostreig va seguir en totes les comissions les indicacions d'aquest tipus de selecció de subjectes. En tots els casos es van triar unes variables per fer els conglomerats i estrats que un cop determinats van servir per produir les mostres convidades.

Cal dir que les mostres resultants i finalment les productores de dades van ser, com es pot apreciar, molt elevades per la qual cosa, des d'aquest punt de vista, la representativitat general estava garantida. A més, en la majoria dels casos les mostres estaven sobredimensionades per a l'estudi general. Aquest és el cas de Catalunya i el País Basc, que van voler seleccionar més subjectes per poder fer, posteriorment a l'estudi estatal, estudis propis amb les dades únicament de la comunitat.

Es pot comentar també que, atesa la complexitat de les qualificacions en expressió escrita i ortografia, les mostres en aquestes dues proves, van ser molt més reduïdes.

#### *Instruments i tècniques d'avaluació*

El quadre 3 recull una síntesi dels instruments i de les tècniques que es van utilitzar en el diagnòstic<sup>7</sup>.

6. Per als lectors interessats en aquesta qüestió, tots els informes publicats, especialment el referit a resultats escolars (MEC: 1998c) recullen apartats referits al mostreig amb més detall del que aquí es presenta de manera sintètica.
7. La comissió IV va encarregar a més informes a un nombre limitat d'experts sobre la formació inicial i permanent del professorat. Van contestar 9 experts, 7 espanyols i 2 internacionals.

**Quadre 2.** Elements metodològics del mostreig en cada comissió especialitzada.

Població	Mostra	Procediment	Variables
I Tots els subjectes escolaritzats en cursos modals de 14 i 16 anys <sup>1</sup> . Excepte Andalusia en l'origen, Canàries en els resultats i el País Basc en els de 16 anys <sup>2</sup>	20.642 (14 a.) 761 centres 25.893 (16 a.) 913 centres. Expr. I Ortog. 1.700 (14 a.) 66 centres 1.680 (16 a.) 64 centres	Estratificació Conglomerats 2 etapes Aleatori simple Error mostral i correlació intraclasse	Autonomia Edat Línia curricular Titularitat Província Modalitat lingüística (Euskadi) Grandària poblacional (Catalunya)
II Professors que el curs 1996-97 impartien classes d'ESO excepte Andalusia, Canàries, Ceuta i Melilla	3.287 docents	3.287 docents Etapes Sorteig probabilístic	Titularitat Mida de la població Comunitat Matèries
III Tots els centres públics i privats concertats en els quals s'impartia, almenys, un curs d'ESO	4.661 persones 565 centres 10 persones per centre, dels diferents estaments	Estratificació Selecció intencional i aleatòria	Titularitat Àmbit territorial Línia curricular
IV Professors que el curs 1996-97 impartien classes d'ESO, excepte Andalusia, Canàries, Ceuta i Melilla	3.026 professors 619 centres	Estratificació Etapes Sorteig probabilístic	Titularitat. Mida de la població Comunitat  Matèries
V Pares o tutors d'alumnes que fessin algun dels cursos modals de 14 i 16 anys	11.481 famílies	Estratificació Conglomerats 2 etapes	Autonomia Edat Línia curricular Titularitat Província Modalitat lingüística (Euskadi) Grandària poblacional (Catalunya)

1. Cursos modals de 14 i 16 anys són aquells en els quals les puntuacions en la variable edat més freqüent (la moda) és 14 i 16 anys. Això aplega 2n d'ESO i 8è d'EGB per als 14 anys i 4t d'ESO, 2 d'FP1 i 2n de BUP per als 16 anys.
2. Andalusia no va participar; Canàries sí, però va decidir no incloure els seus resultats en l'anàlisi de dades. Per tant les conclusions no s'hi podran generalitzar. El País Basc només va participar en el grup d'edat dels 14 anys.



**Quadre 3.** Instruments i tècniques utilitzades.

	<b>Instruments</b>	<b>Tècniques</b>
I	Qüestionaris-prova de rendiment en matemàtiques i en llengua i literatura. A Catalunya, castellà i català	—————
II	Qüestionari	Grups de diagnòstic: Professors: 3-4 professors de Primària i 3-4 professors de Secundària més un moderador Directors: 2-3 directors, 2-3 inspectors i experts més un moderador En total 252 participants més els corresponents moderadors
III	Qüestionari	—————
IV	Qüestionari	Taules rodones: 63 professors Una taula per comunitat autònoma amb competències i dues per al territori MEC
V	Qüestionari	—————

En general, per a mostres tan nombroses, totes les comissions van optar pel qüestionari, en forma de prova de rendiment en la primera comissió i qüestionaris d'enquesta per a les altres.

Dues comissions van optar per complementar aquesta font d'informació amb la utilització de grups de discussió (anomenats per les comissions *grups de diagnòstic i taules rodones*). Aquests grups i taules rodones, amb una durada de 2 hores, servien per poder debatre, complementar i enriquir la resta de dades.

*Elaboració dels instruments*

Per a totes les comissions, en l'elaboració dels instruments van participar els equips de suport als membres de les comissions a cada comunitat autònoma. Aquests equips van fer les propostes d'ítems dels quals es va fer la primera tria per tal d'elaborar les primeres versions dels instruments que serien administrats en una prova pilot abans d'elaborar els instruments definitius.

En la comissió I, les propostes d'ítems van ser triades d'acord amb una matriu d'especificacions en la qual es tenien en compte els continguts i les habilitats requerides (en forma de dominis cognitius), que van donar lloc a les esmentades proves pilot. Els grups pilot tan sols es van seleccionar del terri-

tori MEC<sup>8</sup> per tal de revisar-ne les versions, fer-ne el control de fiabilitat, dificultat, etc., i poder confeccionar així les proves definitives.

La fiabilitat, o estabilitat en la mesura, es va comprovar mitjançant l'estadístic Alpha de Cronbach que en tots els casos va donar valors superiors a .80, motiu pel qual es va considerar que tots els qüestionaris presentaven una fiabilitat ben acceptable.

Respecte a la validesa o capacitat d'un instrument per mesurar allò que vol mesurar, per la participació d'experts de totes les comunitats que van fer propostes i les van revisar, partint a més del currículum de les edats en qüestió que servia de guia, es pot afirmar que la validesa de les proves va ser també acceptable segons els judicis dels experts que van actuar com a jutges<sup>9</sup>.

Els índexs de dificultat i discriminació van servir per fer proves que poguessin incloure ítems de tres tipus amb dificultat variable. Aquesta dificultat va servir també per relacionar els ancoratges de dificultat de l'escala arbitrària teòrica amb els ítems dels qüestionaris.

Mereix un comentari la metodologia emprada per a la confecció de les proves de rendiment i dels criteris per a la seva valoració.

En mesurament i avaluació del rendiment s'utilitzen dues teories:

- La teoria clàssica dels tests (TCC).
- La teoria de resposta a l'ítem (TRI).

La primera s'utilitza tradicionalment i té en compte, per a la interpretació dels resultats, el conjunt de la prova i expressa aquests resultats obtinguts en percentatges d'encerts i d'errors.

La teoria de resposta a l'ítem se centra, en les seves anàlisis, en cadascun dels ítems d'una prova per separat. La unitat ja no és la prova sinó l'ítem. Cada ítem és analitzat separatament posant en relació la capacitat que es vol mesurar amb la probabilitat de respondre correctament l'ítem. Es parteix de la unidimensionalitat<sup>10</sup> de l'escala de rendiment a la qual contribueix cada ítem per separat, de manera que es pot situar a qualsevol subjecte en una escala que mesura capacitats. Es tracta d'un procediment criterial que fixa els

8. Possiblement hagués estat recomanable fer el procés pilot a centres de totes les comunitats per esbrinar el seu funcionament, les dificultats de comprensió, etc., sense que això justifici els resultats finals obtinguts. En ciències humanes, hi ha tantes fonts de variància en els resultats derivats de l'aplicació dels instruments que com més fonts es puguin aïllar o conèixer, més garanties es tenen que els resultats no estan esbiaixats.
9. Una altra qüestió seria si aquell currículum és el més indicat per a aquelles edats i en aquell context socioeconòmic i cultural, en la selecció dels seus continguts, en la seva seqüenciació o, fins i tot, en la manera de ser avaluats (avaluar resultats, processos...). Aquestes qüestions, però, no van formar part del diagnòstic pel fet de no estar previstes ni en el disseny inicial ni en el creuament de dades entre les comissions. Per això d'acord amb la normativa el currículum pot quedar ben representat sense qüestionar-ne els propis continguts.
10. La unidimensionalitat d'una escala es deriva d'acceptar que l'escala només mesura un tret (habilitat o atribut) que explica el comportament dels subjectes en la prova. No sempre les escales són unidimensionals per la qual cosa es pot primer comprovar l'esmentada unidimensionalitat (Del Rincón et al.: 1995). En l'estudi es va partir de la unidimensionalitat.

valors de referència amb independència de les mostres o grups a qui s'apliqui una prova.

La TRI assigna una escala de referència segons la qual es poden interpretar els resultats de qualsevol subjecte que hagi respost a qualsevol prova de rendiment, elaborada amb aquest sistema.

Aquesta escala criterial, que es va fixar com si es tractés d'una prova teòrica de 500 punts, es va dividir en ancoratges que es relacionen amb habilitats de creixent dificultat i complexitat.

### *Característiques de les proves*

En les proves de llengua es van tenir en compte quatre apartats:

- Comprensió lectora.
- Aplicació de regles lingüístiques i literatura.
- Ortografia.
- Expressió escrita.

L'ortografia es va mesurar mitjançant un dictat i l'expressió escrita, amb una redacció.

Les proves eren del tipus objectiu de 40 ítems amb quatre opcions de resposta de les quals només una era correcta. Aquest tipus de prova és molt objectiva i no permet matisos, la qual cosa és també una limitació que impedeix esbrinar quins processos ha seguit l'alumne o fins a quin punt té o no té algun coneixement respecte al resultat que se li demana.

En les proves de matemàtiques es van establir cinc apartats:

- Nombres i operacions.
- Mesura.
- Geometria.
- Anàlisi de dades.
- Àlgebra i funcions.

Els 45 ítems en aquest cas tenien cinc opcions de resposta.

A més de les proves de rendiment es va administrar un instrument amb 46 ítems sobre el context sociofamiliar de l'alumne. Aquest instrument va servir per poder fer creuaments de variables implicades combinant, com és habitual en aquests estudis, l'estricta avaluació del rendiment o millor, dels resultats acadèmics, amb estudis d'investigació sobre l'efecte d'altres variables.

En l'estudi dels resultats acadèmics l'escala per a la interpretació dels resultats de les proves, com ja s'ha dit, tenia un recorregut de 0 a 500 amb ancoratges cada 50 punts. A continuació s'exposa, a tall il·lustratiu, l'assignació d'habilitats als ancoratges de l'escala de llengua, en l'apartat de comprensió lectora (MEC: 1998c) (quadre 4).

Els punts d'ancoratge esmentats proporcionen un criteri per valorar si un subjecte domina una capacitat o no.

**Quadre 4.** Els punts d'ancoratge en l'escala de comprensió lectora.

<b>Nivell</b>	<b>Descripció</b>
150	Comprèn el significat d'enunciats específics de textos utilitaris. Integra informació d'un text per resoldre problemes presentats en textos informatius.
200	Comprèn el significat de paraules específiques, reconeix el tema o idea principal i estableix la seqüència temporal de textos utilitaris.
250	Comprèn el significat de paraules abstractes o altres poc habituals. Comprèn el significat literal (idea central, personatge i tesi de l'autor). Interpreta les idees principals i valora elements retòrics de textos informatius. Integra la informació de textos informatius i literaris per obtenir-ne una de nova i reconeix el doble sentit.
300	Comprèn les idees secundàries i enunciats sintàctics complexos dels textos informatius. Reconeix i interpreta el sentit figurat i el doble sentit de textos informatius i literaris.
350	Comprèn les idees abstractes dels textos informatius.
400	Comprèn les idees abstractes en textos utilitaris i informatius. Reconeix i valora la precisió de les idees i les tesis defensades en textos informatius. Interpreta el significat de procediments retòrics utilitzats en textos literaris.

Encara faltava, un cop determinats els ancoratges i les habilitats que recollien, establir els mínims acceptables a l'escala. La comissió especialitzada I va rebre el vistiplau del comitè científic de l'INCE<sup>11</sup> per situar el punt de tall als catorze anys en l'ancoratge o nivell 250 i el punt de tall als setze en el 300.

Per a la resta de comissions, la feina de confeccionar els instruments va ser molt més senzilla. Es va fer també un primer banc de preguntes amb els corresponents ítems. Dels bancs es van seleccionar els considerats com a més pertinents i amb aquests es va elaborar la prova pilot que també va servir per revisar i elaborar les versions definitives.

En general el nombre de preguntes no va ser massa elevat si bé cal tenir en compte que cada pregunta incloïa un nombre variable d'ítems (entre 4 i 8), per la qual cosa els qüestionaris tenien un cert recorregut. Les opcions de resposta també van ser objecte de consideració.

En els qüestionaris que pregunten sobre l'estimació, percepció o valoració que fa un subjecte sobre algun fenomen és rellevant considerar si proporciona un nombre parell o senar d'opcions de resposta.

El quadre 5 recull les característiques dels instruments utilitzats a la resta de comissions.

11. És un òrgan consultiu de l'INCE integrat per set experts: una presidenta i sis vocals.

**Quadre 5.** Característiques dels instruments de mesura de les comissions II a V.

<b>Comissió</b>	<b>Nre. d'ítems</b>	<b>Tipus</b>	<b>Dimensions</b>
II	56 preguntes amb $n$ ítems cada una	5 opcions de resposta referides a l'acord, importància o freqüència	Intencions, continguts i objectius de l'ESO PCC i programació Mitjans i recursos didàctics Procés d'ensenyament Diversitat Integració d'alumnes amb NNEE Avaluació Clima i cultura escolar Mitjans de suport
III	37 preguntes amb $n$ ítems cada una	5 opcions de resposta referides a l'acord en funció de la situació real i la desitjable	Director Funció directiva Participació Convivència als centres
IV	62 preguntes amb $n$ ítems cada una	4 opcions de resposta referides a la freqüència o a la quantitat de l'estímul	Formació inicial del professorat Planificació i programació Acció a l'aula i al centre Planificació de l'orientació Promoció professional
V	51 preguntes amb $n$ ítems	4 opcions de resposta referides a la freqüència o quantitat de l'estímul	Recursos educatius a la llar Els pares i l'educació dels fills Formació d'actituds i valors Relació escola-família

**Quadre 6.** Distribució de la mostra en els diferents nivells de les diferents proves.

<b>Comprensió lectora</b>				
<b>Rangs</b>	<b>14 anys Percentatge</b>	<b>Acumulat</b>	<b>16 anys Percentatge</b>	<b>Acumulat</b>
101-150	5.30	5.30	0.10	0.10
151-200	25.70	31.00	4.6	4.70
201-250	44.30	75.30	22.3	27.00
251-300	22.60	97.90	50.2	77.20
301-350	2.10	100.00	21.3	98.50
351-400			1.5	100.00
<b>Regles lingüístiques i literatura</b>				
101-150	1.80	1.80		
151-200	24.40	26.10	8.00	8.00
201-250	45.50	71.60	27.70	35.70
251-300	26.40	98.10	40.10	75.80
301-350	1.90	100.00	21.20	97.00
351-400			2.90	99.90
401-450			0.10	100.00
<b>Matemàtiques</b>				
101-150	0.90	0.90	0.10	0.10
151-200	26.90	27.80	10.20	10.30
201-250	44.10	71.80	28.20	38.50
251-300	24.30	96.10	39.10	77.50
301-350	3.90	100.00	19.70	97.20
351-400			2.80	100.00

#### **4. Alguns comentaris als resultats del diagnòstic**

##### *Els resultats acadèmics*

Tal com va ser dissenyat l'estudi, els resultats acadèmics van ser l'aspecte més comentat i més extensament treballat. Per aquest motiu s'inclou una petita mostra dels resultats globals de llengua castellana i de matemàtiques amb alguns comentaris. Cal tenir present, però, que aquests resultats mereixerien un article específic i, per tant, donat el caràcter global d'aquest text, no es comentaran a bastament.

El quadre 6 recull els esmentats resultats utilitzant, com s'ha especificat, l'escala de 500 punts i en relació amb aquesta escala, la distribució, en percentatges, dels alumnes de 14 i 16 anys.

Els resultats, tal com es pot comprovar, se situen, globalment, per sota dels límits o punts de tall considerats com a acceptables. Als 14 anys un grup molt nombrós de subjectes se situa per sota del nivell 250 de referència i als 16 anys passa el mateix amb el nivell 300.

Els resultats, en general, situen les dues poblacions estudiades en el nivell immediatament inferior al considerat mínim o acceptable de referència. Aquesta qüestió cal tenir-la present. D'altra banda, les dades que ha proporcionat aquest estudi no són contradictòries amb les dades d'estudis anteriors, estatals i internacionals.

Quant a l'expressió escrita, s'observa un millor nivell de composició als 16 anys que als 14. També es detecten millores en la redacció i presentació del text, que a més s'incrementa en extensió. De tota manera es constaten problemes en l'expressió que afecten, d'acord amb els nivells considerats satisfactoris, més del 50% dels alumnes.

### *Variables que intervenen en els resultats*

Les variables que intervenen en els resultats són aquelles que no formen part de l'avaluació de manera directa però se sospita que poden explicar una part de la variabilitat dels resultats.

Respecte a aquestes variables estudiades, l'informe avisa que s'han de prendre amb molta cautela i prevenció. D'aquestes, en seleccionem algunes per comentar-les.

Quant al sexe, a l'informe es constaten diferències que estan a favor de les alumnes en alguns casos i a favor dels alumnes en altres. El text, però, deixa entreveure que sembla detectar-se una tendència a superar les diferències en funció de grups i d'edats.

Una altra variable, controvertida per l'ús que se'n pugui fer, és la de les diferències en funció de la titularitat dels centres. L'estudi sembla confirmar que les diferències són significatives a favor dels centres de titularitat privada. També se cita en el text la possible incidència d'altres variables de tipus socioeconòmic provinents de l'entorn familiar que poden explicar la variabilitat detectada. Algunes d'aquestes variables són també controlades en l'estudi per trobar l'anomenada *eficàcia interna*<sup>12</sup> dels centres (equitat) i poder-la contrastar amb l'eficàcia directa, que és la que en principi es detecta. Si bé l'estudi no és concloent, apunta al fet que d'eficàcia interna (un cop controlades les variables que hi intervenen) també en tenen més els centres privats. Aquest últim extrem, que pot haver estat confirmat empíricament, pot justificar dues intervencions diferents com a resultat de l'última fase d'un procés d'avaluació (que esmentàvem com a presa de decisions per a l'acció). En aquest sentit l'estudi, en les seves conclusions, afirma que cal potenciar les possibles deficiències que

12. L'eficàcia interna o equitat és concebuda en l'informe com la capacitat que té un centre per superar les dificultats que provenen de variables socioeconòmiques, culturals, etc., i aconseguir els objectius educatius, «neutralitzant» l'efecte negatiu que aquestes tenen.

l'escola pública pugui tenir i que poden tenir relació amb els resultats esmentats tot i que altres col·lectius poden veure en aquesta constatació un argument de pes per justificar un nou increment de l'impuls, gens legítim en la nostra opinió, cap a la privatització de l'ensenyament obligatori. En aquest punt l'avaluació passa a estar marcada encara més pel caràcter axiològic dels criteris i de les opcions.

Un dels punts més esperats de l'informe va ser el de les diferències en funció de la línia curricular o sistema educatiu (reformat o no reformat). D'alguna manera es pretenia des dels diferents sectors polítics poder utilitzar els resultats per censurar o ratificar la bondat del nou sistema educatiu desenvolupat a partir de la LOGSE. Cal puntualitzar dues qüestions:

- En primer lloc per les característiques de la situació educativa a Espanya no era possible utilitzar l'informe com a argument ja que no a totes les autonomies hi havia el mateix grau d'implantació de la reforma educativa ni fins i tot dins una mateixa autonomia, per la qual cosa les dades, com ja s'ha esmentat, es referien a nivells educatius i currículums diferents.
- En segon lloc, a banda d'haver de disposar d'un mateix tipus de dada, cal disposar de dades en evolució. Per això són importants els estudis diacrònics longitudinals. Proporcionen informació «en moviment» sobre el sistema i el seu procés de consolidació i desenvolupament.

### *Altres resultats rellevants*

Del treball de la resta de comissions, en destaquem els resultats i els comentaris següents:

- Els resultats confirmen algunes de les opinions que informalment es comenten en contextos educatius sobre el sistema educatiu en aquestes edats. Aquesta corroboració empírica matisa també algunes qüestions. Per la representativitat de les mostres de professors i professores dels quals es recull la opinió, aquesta informació és especialment rellevant.
- Contràriament al que alguns podrien pensar, els plantejaments de la reforma semblen un marc legal adequat per a la majoria dels enquestats. Si bé se'n fan moltes crítiques, aquestes es refereixen més a la seva aplicació o bé a millorar-ne aspectes concrets i no pas a canviar de marc.
- Respecte als documents de gestió del centre, es detecta el perill de la seva burocratització. Molts docents afirmen que als seus centres se'n fa escàs o nul ús i que si s'elaboren és degut a la imposició normativa. En general es valoren més del que es fan servir.
- Els docents valoren la flexibilitat com els llibres de text enfoquen els tercers nivells de concreció. D'altra banda, també destaquen que això pot tenir també conseqüències disfuncionals quan, per exemple, es fan trasllats d'alumnes d'uns centres a uns altres.



- Els docents manifesten tenir dèficits en la formació inicial (especialment els de Secundària) i en la formació continuada<sup>13</sup>.
- Es constata una visió positiva que tenen els docents d'ells mateixos, com també succeeix (en major grau fins i tot) amb els càrrecs directius i amb els pares i mares. Tot i això es constata una percepció de ser mal valorats per l'opinió pública alhora que es detecta una crisi d'identitat per a les noves funcions i encàrrecs que la societat sembla abocar a l'escola.
- La diversitat a l'ESO i les implicacions que aquesta té en la concreció del currículum i en la gestió dels processos a l'aula és un altre aspecte que preocupa els docents que se senten insegurs, sobretot per venir d'un sistema en el qual l'homogeneïtat era elevada.
- Indefinició dels nous serveis d'orientació i d'assessorament intern i extern als centres. L'orientació, els plans d'acció tutorial i la figura del psicopedagog són encara qüestions que no estan prou perfilades, segons se'n desprèn de l'estudi.
- S'aprecia una gran distància entre la valoració i l'aplicació dels diferents mitjans i recursos didàctics a l'abast dels docents. En general es valoren molt alguns recursos, però només s'utilitzen els més tradicionals.
- Es detecten problemes en la nova concepció de l'avaluació i promoció dels alumnes que planteja nombrosos dubtes i problemes als docents. En general es valoren més que no pas s'utilitzen els procediments avaluatius menys convencionals (els exàmens escrits es valoren menys que s'utilitzen i l'observació de treballs a la inversa).
- La direcció és valorada positivament per tots els enquestats, detectant-se dos grans enfocaments directius: la direcció amb les funcions que li atribueix la normativa, la direcció centrada en el professorat.
- Tampoc es considera que la participació i la implicació als centres estigui en el seu millor moment, com tampoc ho està la participació dels pares i de les mares. Aquesta qüestió de la participació de les famílies és un element rellevant que si es corregís ajudaria notablement a allunyar el presagi de la violència i altres problemes de relació que es comencen a insinuar en l'estudi i en altres treballs similars. Els docents i els càrrecs directius afirmen, però, que el seu funcionament és satisfactori.

## 5. Conclusions i perspectives

L'avaluació del sistema és un procés que, com a tal, ha de desenvolupar-se en contextos d'independència política si bé cadascuna de les decisions que es prenen tenen una justificació que depassa allò tècnic i entra de ple en l'axiologia. Decidir quin serà el disseny d'una avaluació implica prendre una opció i eli-

13. En el moment de fer l'estudi la formació inicial del professorat de Secundària es trobava, també, en procés de transformació sense que es pogués avaluar conclouentment si el nou sistema ha representat una millora respecte a l'anterior, com a mínim en la percepció dels docents.

minar-ne d'altres. El mateix passa amb la selecció dels objectes d'avaluació, els instruments, els criteris per valorar i, finalment, el comentari als resultats obtinguts. De manera que cal destriar el que són les característiques tècniques o metodològiques d'un estudi de la concepció global que inspira el seu disseny. L'estudi de l'INCE, tal com han afirmat molts experts és un estudi seriós que ha pres, però, una opció metodològica concreta a partir d'un disseny fragmentat dels diferents àmbits que s'integren en el sistema educatiu. L'alternativa podria haver estat la vinculació d'àmbits: resultats —currículum—, demandes professionals, etc. Tot i això, cal pensar seriosament en els resultats obtinguts i buscar estratègies per millorar-los en futures ocasions.

Respecte a l'informe i el seu contingut, es pot concloure que els resultats han anat confirmant una tendència ja detectada en anteriors estudis que evidencien que el rendiment no és molt satisfactori, sense que això vulgui dir que sigui desastrós. Els nivells detectats són força similars a tot l'Estat dintre de la franja immediatament inferior als mínims establerts pels experts. L'explotació mediàtica de les diferències en poc ajuden a reflexionar i a actuar amb voluntat de millora.

A banda dels resultats educatius, les altres comissions han anat corroborant i desmentint alguns tòpics sobre el sistema que ja s'han anat comentant en apartats anteriors.

Per les característiques del disseny que s'ha fet servir és difícil iniciar processos de millora quan els resultats estan desvinculats del currículum, de les estratègies, les demandes de l'entorn, la seva implicació, etc. Tot i això cal aprofitar-los perquè sí que proporcionen informació suficient com per justificar algunes accions.

Les perspectives de l'avaluació del sistema podrien recollir les consideracions abans esmentades. Cal millorar els dissenys d'avaluació i cal complementar les avaluacions fetes en diferents moments per obtenir visions de desenvolupament del sistema.

La participació és també un principi bàsic per desenvolupar l'avaluació i en aquest sentit és interessant incloure referències i pistes d'intervenció en tot informe d'avaluació. És recomanable, doncs, pensar en la difusió i utilització dels informes i les seves pautes per part de tots els implicats. És important, doncs, difondre oportunament l'informe i buscar la manera de fer-lo útil per als centres com també buscar maneres de fer avaluacions més rendibles per a l'educació del país.

Un element clau de l'avaluació és el seu caràcter axiològic. Avaluar implica valorar, determinar el valor d'un procés o fenomen educatiu. Quan es posa en joc una activitat valorativa sorgeixen immediatament un conjunt de problemes que fan referència a la jerarquia de valors adoptats, a les persones que sustenten la responsabilitat d'avaluar, als objectius que pretén satisfer l'avaluació i a la dimensió política de tot el projecte d'avaluació. Afegint a aquesta característica la qüestió de la presa de decisions és especialment important atendre a les decisions enfocades al millorament del sistema que prendran els diferents estaments implicats en aquesta qüestió: l'administració en els seus diferents nivells i també

els professionals i tots els agents educatius que, dia a dia, fan possibles uns resultats o uns altres. L'avaluació entesa d'aquesta manera no garanteix el canvi i la millora, però la possibilita i estimula. És per això que es considera molt important construir i millorar la cultura avaluativa en l'educació.

Aquesta visió del fet avaluatiu, si bé complexa i controvertida, ha d'instal·lar-se permanentment en el quefer de tots els professionals i agents vinculats a l'educació en tots els seus nivells per aconseguir una pràctica contínua, participativa i crítica de la millora de l'educació.

## Bibliografia

- BISQUERRA, R.; MARTÍNEZ, M. (1998). *El clima escolar als centres d'ensenyament secundari a Catalunya*. Informes d'avaluació núm. 1. Barcelona: Consell Superior d'Avaluació del Sistema Educatiu, Departament d'Ensenyament.
- COMITÉ CONJUNTO DE ESTÁNDARES PARA LA EVALUACIÓN EDUCATIVA (1997). *Estándares para la evaluación de programas*. Bilbao: Ediciones Mensajero.
- COOMBS, P.H. (1985). *La crisis mundial de la educación. Perspectivas actuales*. Madrid: Aula XXI-Santillana.
- Decret 191/1996 pel qual s'aprova el reglament orgànic de centres públics de Catalunya. DOGC 2218, 14-6-96.
- DE LA ORDEN, A. (1985). *Investigación educativa. Diccionario de ciencias de la educación*. Madrid: Anaya.
- DEL RINCÓN, D. et alt. (1995). *Técnicas de investigación en ciencias sociales*. Madrid: Dykinson.
- FERRER, F. (1988). *L'avaluació del sistema educatiu. Aproximació teòrica des de i per a Catalunya*. Barcelona: Consell Escolar de Catalunya (en premsa).
- GENERALITAT DE CATALUNYA (1991). *Avaluar per innovar*. Barcelona: Departament d'Ensenyament.
- (1997). *Sistema d'indicadors d'ensenyament de Catalunya*. Barcelona: Consell Superior d'Avaluació del Sistema Educatiu.
- (1998a). *Avaluació interna de centres*. Barcelona: Departament d'Ensenyament.
- (1998b). *Avaluació interna de centres. Exemples de prova cycle mitjà. Llengua catalana i literatura. Llengua castellana i literatura*. Barcelona: Departament d'Ensenyament.
- GIL ESCUDERO, G. *El proyecto internacional para la producción de indicadores de resultados educativos de los alumnos*. (Proyecto PISA) de l'OCDE [en línia]. INCE. Ministerio de Educación y Cultura. 1998. <http://WWW.INCE.SEE.MEC.ES>. [Data de consulta: 14-7-98].
- GIL FERNÁNDEZ, P. (1983). *Diagnóstico. Diagnóstico pedagógico. Diccionario de las ciencias de la educación*. Madrid: Santillana.
- GOBIERNO VASCO (1997). *Análisis de la situación internacional de la gestión de la calidad en centros de enseñanza*. Document policopiat.
- Ley Orgánica de Ordenación General del Sistema Educativo. BOE 238, 4-10-90.
- Ley Orgánica de Participación, Evaluación y Gobierno de los Centros Docentes. BOE 278, 21-11-95.
- MINISTERIO DE EDUCACIÓN Y CIENCIA (1991). *Plan de evaluación de centros docentes (niveles no universitarios). Programa piloto. Curso 1991-1992*. Madrid: Secretaria de Estado de Educación-Inspección de Educación-CIDE.

- (1997a). *Tercer estudio internacional de matemáticas y ciencias (TIMSS)*. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1997b). *Evaluación de la educación primaria*. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1998a). *Diagnóstico general del sistema educativo. Avance de resultados*. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1998b). *1. Elementos para un diagnóstico del Sistema Educativo Español. Informe global*. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1998c). *2. Los resultados escolares*. Diagnóstico general del sistema educativo. La escuela secundaria obligatoria. 1997. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1998d). *3. Planes de estudio y métodos de enseñanza*. Diagnóstico general del sistema educativo. La escuela secundaria obligatoria. 1997. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1998e). *4. Funcionamiento de los centros*. Diagnóstico general del sistema educativo. La escuela secundaria obligatoria. 1997. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1998f). *5. La profesión docente*. Diagnóstico general del sistema educativo. La escuela secundaria obligatoria. 1997. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1998g). *6. Familia y escuela*. Diagnóstico general del sistema educativo. La escuela secundaria obligatoria. 1997. Madrid: Instituto Nacional de Calidad y Evaluación.
  - (1998h). *Diagnóstico del Sistema Educativo 1997 en los medios de comunicación*. Madrid: Reprografía MEC, Servicios Centrales.
- MATEO, J. (1994). *L'avaluació educativa*. Document policopiat d'ús intern.
- OCDE. (1995). *Education at a glance. OECD indicators*. Paris: OECD.
- OCDE. New developments on education, Employment, Labour and Social Affairs. [En línia]. Capítols CDEF. OCDE. 1998. <http://www.OECD.ORG>. [Data de consulta: 10-7-98].
- UNESCO (1996). *La educación encierra un tesoro*. Informe a la UNESCO de la Comisión Internacional sobre la Educación para el Siglo XXI, presidida per Jaques Delors. Madrid: Santillana.