

L'AVALUACIÓ DE L'ENSENYAMENT DE LES CIÈNCIES

Anna Maria Geli de Ciurana

Dpt. de Dtca. Matemàtiques i Ciències U.A.B.

RESUMEN

En las páginas siguientes se revisan algunos de los programas de Evaluación de la Enseñanza de las Ciencias que han tenido mayor incidencia en la evolución histórica de esta área de conocimiento. Hasta los años 70, el enfoque de la evolución de las ciencias era, esencialmente, memorístico. A partir de los años 80 surgen propuestas más interesantes, que plantean con especial interés la evaluación de las actividades prácticas.

ABSTRACT

This paper offers a review of Science assessment activity during the last years. Until the late 1970s or thereabouts, the major emphasis in most cognitive assessment activity in science was the recall of knowledge. In the 1980s more interesting developments are underway, with greater attention than before being paid to the assessment of intellectual and practical skills and abilities.

La situació de l'avaluació de les ciències

«Els models d'ensenyament actuals assignen a l'avaluació la funció de control, anàlisi i valoració dels processos d'ensenyament i dels resultats dels programes educatius». (GIMENO SACRISTÁN, 1984). Per assumir aquesta funció, molt més amplia que la clàssica qualificació dels alumnes, la investigació educativa ha dedicat importants esforços al camp de l'avaluació. Des dels anys seixanta s'han desenvolupat diferents models teòrics d'avaluació, que ofereixen alternatives diverses i, a vegades complementàries, per valorar la qualitat dels programes d'ensenyament (STUFFLEBEAM I SHINKFIELD, 1985).

L'aplicació d'aquests models teòrics a la pràctica educativa exigeix un treball de recerca en les didàctiques específiques de cada àrea de coneixements, a les quals correspon de seleccionar els models més adequats al seu propi àmbit i elaborar els instruments necessaris per avaluar els programes de les corresponents matèries.

L'avaluació de l'ensenyament de les ciències com a objecte d'investigació i estudi té una història molt curta. En la dècada dels anys 60, a causa de la preocupació social que va despertar la ciència i d'un major interès per la recerca educativa, es van iniciar, en diversos països, programes d'avaluació dels coneixements de ciències en els nivells d'ensenyament primari i secundari.

A les pàgines següents farem un repàs de l'avaluació de l'ensenyament de les ciències en els darrers anys, i revisarem alguns dels grans projectes que són punt de partida per a l'estudi d'aquesta avaluació.

La història de l'avaluació de les ciències

Els sistemes d'avaluació que s'han aplicat al llarg de la història de l'ensenyament de les ciències estan definits per dos factors que han evolucionat i canviat molt en els últims anys:

1. El concepte d'avaluació, que ha ampliat i aprofundit el seug significat des de la tradicional avaluació dels alumnes fins a l'actual avaluació de programes.

2. Els models d'ensenyament de les ciències, que, d'acord amb les teories psicològiques i pedagògiques de cada època i responent a diferents conceptes de ciència, han passat per successives etapes: l'ensenyament tradicional, entès com a simple transmissió d'informació, l'ensenyament pel descobriment guiat, d'influència conductista, l'ensenyament basat en l'estudi de l'entorn,... fins arribar als actuals models constructivistes.

La combinació d'aquests dos factors és la causa decisiva dels canvis que experimenten els models d'avaluació.

El concepte d'avaluació

STUFFLEBEAM recull el significat vulgar del terme avaluació i el defineix com l'estimació sistemàtica del valor o el mèrit d'un objecte (STUFFLEBEAM, 1987).

Per dur a terme aquesta estimació, en el procés d'avaluació educativa es diferencien tres etapes:

1. Recollida d'informació
2. Anàlisi de les dades recollides
3. Judici sobre el valor o mèrit de l'objecte d'avaluació

En el concepte clàssic, l'objecte de l'avaluació és el rendiment acadèmic dels alumnes. La seva finalitat es redueix a la mesura dels coneixements adquirits al final del programa educatiu i a l'emissió, per a cada estudiant, d'un judici de valor en forma de qualificació (POPHAM, 1975).

Aquest concepte d'avaluació, centrat en l'alumne, ha modificat la seva funció de simple informació final dels resultats fins a un procés d'avaluació contínua, de caràcter diagnòstic, que es proposa d'identificar les dificultats dels alumnes i aplicar, en cada cas, el tractament pedagògic més convenient (C. COLL, 1987).

En el concepte actual, l'objecte de l'avaluació educativa no es limita a l'anàlisi del rendiment acadèmic dels alumnes, sinó que també es proposa conèixer la qualitat del programa d'ensenyament. Per avaluar un programa cal recollir informació dels diferents factors que intervenen en el procés docent i del grau en què s'han assolit les intencions educatives del projecte. (PÉREZ GÓMEZ, 1983) (La finalitat d'aquesta informació és orientar sobre les modificacions que cal introduir en el programa per millorar els resultats. (Utilitzem aquí el terme programa com a sinònim de tota acció pedagògica planificada amb el propòsit de produir un aprenentatge; ens referim tant als grans projectes educatius, com a la planificació a curt termini del treball a l'aula.

Els models d'ensenyament de les ciències

En els darrers anys, l'evolució del concepte de ciència i els progressos que s'han produït en el camp de la pedagogia, la psicologia i l'epistemologia han generat diferents maneres d'entendre l'ensenyament de les ciències. Cada model d'ensenyament necessita de sistemes d'avaluació que s'ajustin a les seves intencions educatives.

Durant la primera meitat del nostre segle, l'ensenyament tradicional de les ciències volia aconseguir que l'alumne adquirís una àmplia informació de continguts científics. L'avaluació es reduïa a mesurar la quantitat de coneixements que l'alumne retenia a la seva memòria; aquesta mesura s'expressava per mitjà d'un valor numèric que venia a representar el percentatge

de coneixements que l'estudiant recordava en relació amb tots els que s'havien donat a classe o que podia trobar en el llibre de text.

En els anys sexanta les teories conductistes de l'aprenentatge orienten l'ensenyament vers els models didàctics de caràcter tecnològic, que defineixen el procés d'ensenyament a partir d'una acurada descripció dels objectius educatius. L'avaluació està centrada en comprovar el grau en què l'alumne assolix aquests objectius. A aquesta època corresponen els primers models teòrics d'avaluació que tenen el seu primer antecedent en R.W. TYLER (1934). Els models d'avaluació referits a objectius formulats com a comportaments mesurables són nombrosos i variats (TYLER, 1934; METFESSEL i MICHAEL, 1967); MAGER, 1972; POPHAM, 1975). La preocupació per l'objectivitat els porta a crear una metodologia d'avaluació pròpia, de la qual, la tècnica més representativa i que més influència ha tingut en els models d'avaluació posteriors ha estat el test objectiu.

Aquesta metodologia d'avaluació positivista i quantitativa, es manifesta adequada per a la valoració dels aprenentatges de l'àmbit cognoscitiu, però, molt aviat, es demostra insuficient per a l'avaluació dels aprenentatges dels àmbits psicomotor i afectiu.

En aquesta època s'inicien els primers programes d'avaluació de l'ensenyament de les ciències en diferents països; programes on s'aplica la tècnica del test objectiu per avaluar els nivells cognoscitius de la taxonomia de B. BLOOM (coneixement, comprensió, aplicació, anàlisi, síntesi i avaluació). Els models d'ensenyament de les ciències d'aquests anys es basen en l'aplicació del mètode científic a l'ensenyament i consideren fonamental per a la formació científica, l'aprenentatge de processos i tècniques experimentals, que no poden mesurar-se únicament amb proves de llapis i paper. Això posa de manifest la necessitat d'elaborar nous models d'avaluació que ofereixin tècniques més obertes i flexibles per a la valoració del treball pràctic.

En avançar la investigació educativa, els mateixos autors que en els seus primers treballs seguien el model tylerià i es limitaven a estudiar el rendiment acadèmic dels alumnes mitjançant tècniques d'anàlisi estadística, van ampliant el camp de l'avaluació vers altres aspectes del projecte educatiu. El model d'STUFFLEBEAM (CIPP, 1971) avalua quatre àmbits diferents: context, projecte, procés i producte. Els treballs d'SCRIVEN i STAKE, que l'any 1967 eren clarament quantitativs, també evolucionen i l'any 1975 STAKE proposa l'avaluació com una resposta a les qüestions que es plantegen sobre el programa educatiu (Evaluation responsive) i SCRIVEN, el 1976, presenta l'avaluació sense referència a objectius.

La recerca avaluativa dels anys 70 i 80 porta a un concepte qualitatiu de l'avaluació que aplica tècniques basades en l'observació per descriure els aprenentatges i proposa l'anàlisi de tots els factors que intervenen en el projecte docent. Els models qualitativs més interessants són: l'avaluació il·luminati-

va de PARLETT i HAMILTON (1972), l'avaluació basada en la crítica artística d'ELLIOT EISNER i l'escola de la Universitat d'Stanford (1977), l'avaluació democràtica de McDONALD (1976), i les propostes de CRONBACH (1980), J. ELLIOT (1982),...

Aquests models qualitius d'avaluació responen molt bé als plantejaments actuals de millora de l'ensenyament, basats en un procés continu d'innovació didàctica que confereix a la tasca del professor una dimensió d'investigació a l'aula (STENHOUSE, 1987).

Al mateix temps, aquestes propostes ofereixen als models d'ensenyament de les ciències dels anys 80 millors perspectives d'avaluació en els tres àmbits considerats, actualment, com a fonamentals per a l'aprenentatge de les ciències: els conceptes, les tècniques i les actituds. L'avaluació qualitativa, basada en la descripció, està especialment indicada per avaluar les tècniques i habilitats i, a més, obre camins per identificar les actituds, normes i valors.

L'aplicació en la pràctica educativa de models d'avaluació amplis, que combinin el rigor i precisió de les tècniques quantitatives d'avaluació i la flexibilitat i la riquesa de les tècniques qualitatives, creiem que permetrà en el futur una avaluació completa dels diferents aspectes de l'ensenyament de les ciències.

Els programes d'avaluació de l'ensenyament de les ciències

En estudiar amb atenció els programes d'avaluació de l'àrea de Ciències podem veure que els dos factors que defineixen cada programa són: Els coneixements que el projecte vol avaluar, i les tècniques que utilitza per a l'avaluació.

Els coneixements que avalua cada projecte són els que els autors del projecte consideren que han de tenir els estudiants de ciències als quals s'adreça el programa (com és lògic, de la definició d'aquests coneixements depenen les categories d'avaluació). Els ràpids progressos que es produeixen en la investigació de l'ensenyament de les ciències són la causa que les categories d'avaluació definides fa pocs anys, en el moment present hagin quedat desfasades, i que en els programes actuals la delimitació del que es vol avaluar sigui una tasca laboriosa i polèmica, sotmesa a constants revisions.

Les tècniques d'avaluació que s'utilitzen en cada programa estan condicionades pels tipus de coneixements que s'han d'avaluar i pels models que ofereix en cada moment el camp de la Teoria de l'Avaluació.

Els projectes d'avaluació que analitzarem es poden classificar en dues categories: els projectes d'avaluació externa, realitzats per especialistes en avaluació de forma independent del programa d'ensenyament; i els projectes

d'avaluació interna, aplicats pel professor de la matèria com a un element més del programa d'ensenyament.

Programes d'avaluació externa

L'avaluació externa té dues finalitats:

1. *Contrastar els resultats d'un o més projectes docents.* En aquest cas la tasca s'encomana a especialistes en avaluació que dissenyen el programa de recollida d'informació, l'anàlisi de les dades obtingudes i l'emissió d'un judici sobre el programa avaluat. Aquest tipus d'avaluació és freqüent a EE.UU. però en el nostre sistema educatiu només es realitza per encàrrec de l'administració (actualment la Generalitat de Catalunya està realitzant l'avaluació dels programes experimentals d'ensenyament secundari).

2. *Garantir un nivell de coneixements equivalent entre els alumnes que finalitzen una etapa educativa.* Aquesta avaluació no la porten a terme les mateixes persones que imparteixen l'ensenyament. Un exemple és el sistema educatiu de Gran Bretanya, que controla els coneixements dels alumnes en acabar l'ensenyament secundari; en el nostre país s'havia aplicat en les revàlides de l'antic batxillerat i actualment es pot comparar amb la prova de selectivitat que s'exigeix a tots els alumnes que volen entrar a la universitat.

La influència que les proves comunes tenen en els programes d'ensenyament del lloc on s'apliquen és molt gran: es pot dir que aquestes proves condicionen el model d'ensenyament de tots els centres. Si l'elaboració de les proves té en compte tots els aspectes de l'aprenentatge de cada matèria els resultats poden ser molt positius, però en el cas de l'ensenyament de les ciències és molt difícil que l'avaluació pugui realitzar-se exclusivament amb proves de llapis i paper. A Gran Bretanya s'estan aplicant models nous d'avaluació que inclouen la valoració del treball pràctic, al llarg del curs, en el propi centre.

El primer gran projecte d'avaluació a nivell internacional, i fins ara l'únic, aplicat a dinou països de tot el món, és el que va iniciar en els anys 60 l'Associació Internacional per l'Avaluació del Rendiment Escolar (IEA). El sondeig s'adreçava a alumnes de 9, 13 i 17 anys i, en una primera fase, es va aplicar fins l'any 1975. Els instruments utilitzats eren tests objectius (els mateixos a tots els països) i s'avaluaven diferents matèries: matemàtiques, ciències, llengua materna, idioma estranger, ... Al mateix temps recollen informació, mitjançant enquestes, sobre les característiques personals, escolars, socials i familiars de cada alumne (COMBER i KEEVES, 1973).

En els anys en què es va aplicar aquest programa d'avaluació les teories conductistes de l'aprenentatge estaven en plena vigència i les taxonomies d'ob-

jectius educatius definien els programes d'ensenyament i la seva avaluació. Els coneixements de ciències que es van mesurar en els tests es definien en l'àmbit d'objectius cognitius de la Taxonomia de B. BLOOM, especialment: Coneixement, Comprensió, Aplicació i Avaluació.

La tècnica de tests objectius que es va aplicar en aquest projecte és, sens dubte, el tipus de prova més adequat per avaluar milions d'estudiants i establir comparacions entre els diversos factors del sistema educatiu. Una mostra tan àmplia només ha estat possible en l'època de la informàtica i dels programes d'estadística per ordinador. No obstant això, les delimitacions dels tests objectius es manifesten especialment en l'avaluació dels coneixements de ciències, en comprovar que no poden aplicar-se a l'avaluació del treball pràctic de l'alumne en el laboratori. Precisament fa vint anys era el moment en què els models d'ensenyament de les ciències proposaven un ensenyament basat en l'experimentació i valoraven especialment l'aprenentatge de tècniques i processos. En aquella època, però, els models d'avaluació no oferien tècniques ni instruments adequats per valorar els coneixements pràctics i només dos països van aplicar proves d'avaluació de les pràctiques, de manera molt restringida.

L'informe dels resultats obtinguts es basa en l'anàlisi estadística de les correlacions entre els coneixements de ciències i l'edat, el sexe, el medi familiar i el país.

A partir de 1984 s'han realitzat nous sondeigs, aplicant les mateixes tècniques d'avaluació de caràcter quantitatiu, i els resultats obtinguts s'han comparat amb els anteriors. Els resultats de les dues sèries d'avaluacions són, bàsicament, coincidents (S. JOHNSON, 1987).

El Programa Nacional d'Avaluació del Progrés Educatiu dels EE.UU. (NAEP) és un altre programa de gran abast (inspirat en la IEA), que es proposa de controlar regularment el rendiment escolar dels estudiants de 9, 13 i 17 anys i d'adults joves dels Estats Units.

A les seves primeres edicions (NAEP, 1975) els objectus del programa eren:

- Conèixer els aspectes fonamentals de la ciència.
- Comprendre els coneixements científics i aplicar-los a problemes diversos.
- Apreciar els coneixements i processos de la Ciència, les seves conseqüències, els límits i la importància que té, juntament amb la tecnologia, en la vida de les persones i de la societat.

En successives edicions el programa ha anat ampliant i aprofundint aquests objectius, fins que en les últimes edicions, es proposa avaluar:

- Continguts científics.

- Capacitat d'investigació (Mètode científic).
- Ciència - Tecnologia - Societat.
- Actituds.

Els coneixements que s'avaluen han evolucionat des de la taxonomia de BLOOM fins a les actuals propostes constructivistes.

Els instruments utilitzats segueixen essent els tests de preguntes de resposta unívoca, però el plantejament de les preguntes fuig del marc escolar i està orientat a situacions quotidianes dels estudiants: interpretació de gràfics, capacitat de deducció, etc...

A partir de 1975 s'inicià a Bèlgica un Projecte Nacional d'Avaluació del Progrés Educatiu, realitzat per la Universitat de Liege, amb la participació de nombrosos pedagogs De Ladsheere, Grisay, Henry, Deltour,... Fins l'any 1979 va estar adreçat a l'ensenyament primari i des d'aleshores fins el 1984 a l'ensenyament secundari (DELTOUR, 1984).

A l'àrea de Ciències els objectius que s'han d'avaluar estan classificats en vuit categories:

- Observació.
- Emissió d'hipòtesi.
- Completar l'observació amb documentació.
- Experimentar.
- Interferir, verificar i predir.
- Aplicar i transferir.
- Mesurar.
- Desenvolupament del sentit pràctic.

Per avaluar aquests objectius han elaborat dues-cents trenta qüestions objectives, la majoria de les quals són les característiques preguntes d'opció múltiple i d'ordenació. El tractament estadístic dels resultats estudia la dificultat de cada pregunta i la fiabilitat dels deu tests en què s'han agrupat.

També han estudiat, mitjançant l'ordinador, els resultats de cadascun dels vuit objectius analitzats.

Al mateix temps s'han passat enquestes als directors dels centres, als professors i als alumnes per conèixer la influència del medi escolar en l'aprenentatge dels alumnes. Mitjançant el càlcul de correlacions i de regressió de variables s'ha intentat identificar els factors que influeixen de manera significativa en l'aprenentatge de les ciències.

A Gran Bretanya, a partir de 1980, s'han desenvolupat dos projectes d'avaluació de l'ensenyament de les ciències que han tingut una forta repercussió en els programes escolars de tot el país. Són: Assessment Performance Units (APU) desenvolupat a Anglaterra, País de Gales i Irlanda del Nord, i Tech-

niques for the Assessment of Practical Skills in Foundation Science (TAPS) a Escòcia.

El projecte APU realitza un sondeig anual dels coneixements de ciències dels alumnes d'11, 13 i 15 anys. En el disseny d'aquest programa, realitzat amb rigor i profunditat, han participat les Universitats de Londres i Leeds, especialment els professors R. GOTT, S. JOHNSON i P. MURPHY (APU, 1985).

L'APU parteix del concepte que defineix la Ciència com a mètode per resoldre problemes i planteja als alumnes activitats classificades en sis categories d'avaluació:

1. Ús de gràfics i representacions simbòliques
2. Ús d'aparells i instruments de mesura
3. Observació
4. Interpretació i aplicació
5. Planejar investigacions
6. Realitzar investigacions

Les tècniques d'avaluació que s'apliquen per mesurar les categories 1, 4 i 5 són proves escrites. En alguns casos les preguntes són tancades, però sovint demanen a l'alumne que justifiqui la seva resposta. Aquest plantejament permet a l'estudiant expressar-se lliurement i recollir les observacions com en una pregunta oberta (HARLEN, 1981; DRIVER, 1982).

Per mesurar les categories 2 i 3 es fan proves experimentals que consisteixen en oferir a l'alumne el material necessari per dur a terme una determinada experiència i a continuació es plantegen qüestions molt precises sobre l'activitat realitzada. Les experiències es realitzen en grups reduïts. (WELFORD, 1985; GOTT i MURPHY, 1987).

Per avaluar la capacitat de realitzar investigacions cal que cada alumne, de manera individual, desenvolupi una petita investigació que pot durar entre 45 i 60 minuts per als alumnes de primària i 30 minuts per als de secundària. Aquest sistema d'avaluació, l'únic que pot donar informació sobre la capacitat de recerca de l'estudiant, té l'inconvenient que és molt lent i no permet avaluar més de 5 o 6 alumnes per dia.

L'avaluació del treball pràctic de ciències és especialment difícil perquè les tècniques de recollida de la informació han de reflectir tota l'actuació de l'alumne i això només s'aconsegueix amb tests pràctics del tipus que proposa l'APU: proves que combinen l'observació del treball de l'alumne amb qüestions escrites sobre la interpretació que fa de l'experiència.

Per poder establir comparacions entre els resultats de l'avaluació del treball pràctic cal que es realitzi en circumstàncies equivalents i que els professors que apliquen les proves estiguin preparats per realitzar la seva interpretació. Per garantir la preparació del professorat que participa en el

programa, tots els avaluadors segueixen cursos a les Universitats de Londres o Leeds. Utilitzen el video per gravar les sessions de treball en grup i disposen de llistes de control de les activitats. A fi que les circumstàncies d'aplicació siguin equivalents es faciliten a les escoles tots els materials i aparells necessaris.

A partir de 1988 s'ha iniciat un nou programa APU d'avaluació dels coneixements de tecnologia, adreçat a estudiants de 15 anys. Aquest projecte ha despertat l'interés dels educadors perquè es proposa seguir, per primera vegada en l'àrea de ciències, el model d'avaluació il·luminativa de PARLETT i HAMILTON. Aquest model defineix l'avaluació com a una visió comprensiva de tot el medi docent, i considera que els dos factors claus per a l'avaluació d'un programa són: els sistema d'ensenyament i l'ambient d'aprenentatge (PARLETT i HAMILTON, 1972). Per obtenir la informació pròpia utilitzar l'observació, les entrevistes amb totes les persones implicades en el programa d'ensenyament (professors i alumnes), enquestes, tests i la consulta de fonts documentals i històriques. D'aquest projecte no disposem, encara, de resultats publicats.

El TAPS (BRYCE et al, 1983) es un programa d'avaluació del treball pràctic de Ciències que ha definit les següents categories d'avaluació:

- Habilitats per observar.
- Habilitats per recordar.
- Habilitats per seguir instruccions.
- Habilitats per mesurar.
- Habilitats de manipulació.
- Habilitats d'inferència.
- Habilitats processuals.
- Habilitats per seleccionar processos.

Les tres últimes categories s'han inclòs en les últimes edicions del programa. (BRYCE, 1987).

Per a cada objectiu es presenta a l'alumne la situació real d'una pràctica de laboratori i se li plantegen diverses qüestions tancades i obertes a les quals pugui desenvolupar les corresponents habilitats.

Uns altres programes d'avaluació externa són els aplicats institucionalment amb el propòsit de garantir un nivell homogeni de coneixements en els alumnes que acaben un determinat nivell d'ensenyament.

A Gran Bretanya en acabar l'ensenyament secundari s'apliquen proves d'avaluació comunes a tots els centres de cada districte educatiu. Aquests exàmens externs són obligatoris per a tots els alumnes que volen obtenir el Certificat d'Educació Secundària (GCSE).

Des de 1986 s'aplica en aquestes proves un nou sistema de valoració que

inclou a la qualificació, l'avaluació del treball pràctic. El nou disseny d'aquestes proves en l'àrea de ciències ha estat molt influït pels dos programes que acabem d'esmentar: l'APU i el TAPS.

Les proves de ciències tenen com a objectiu comprovar:

- La comprensió de les matèries bàsiques.
- La interpretació de les dades científiques.
- La investigació.

Les tècniques analitzades són l'observació, la mesura, la interpretació de resultats i el treball en equip.

Al final de l'etapa de secundària es fan proves escrites que, a cada districte educatiu, són elaborades per una comissió d'especialistes que dissenyen els exàmens amb dos anys d'antelació. L'avaluació del treball pràctic la realitza, al llarg del curs, el professor de cada centre. La seva qualificació representa un 25 % de la nota final. (Northern Examining Association, 1988; BBC Scotland, 1987).

L'anàlisi de les proves de selectivitat realitzades en l'àrea de ciències, a les universitats espanyoles demostren que es demana als alumnes un tipus de coneixements molt restringits, sovint memorístics, que no es corresponen, ni de lluny, als aprenentatges que proposen els models actuals d'ensenyament de les ciències (AGUILA, GIL i GONZÁLEZ, 1988). Amb aquesta prova no es controla la qualitat de l'ensenyament de les ciències.

Programes d'avaluació interna

L'avaluació planificada pel propi professor, com a un element més del projecte docent, necessita estar ben dissenyada per complir els seus objectius:

1. *Ajustar la planificació de l'ensenyament a les característiques individuals dels alumnes.* Segons el moment en què es realitza pot ser inicial o contínua.

La importància de l'avaluació inicial en l'ensenyament de les ciències ha estat àmpliament demostrada en els nombrosos estudis realitzats sobre errors conceptuals dels estudiants. En ells es posa de manifestat la incidència dels coneixements previs en l'aprenentatge de nous conceptes (AUSUBEL, 1978; HILL, 1985; HEWSON, 1982; GIORDAN, 1985).

L'avaluació contínua, al llarg del programa, permet detectar les dificultats de cada alumne, i intervenir amb l'ajuda pedagògica més convenient per a la seva formació (avaluació formativa) (SCRIVEN, 1967). Al mateix temps analitza el desenvolupament del programa.

2. *Determinar, al final del programa, el grau en què s'han assolit les seves intencions.* Cada projecte d'ensenyament té una intencionalitat que queda definida en el programa. L'avaluació dels resultats obtinguts, realitzada en l'avaluació final o sumativa, informa de la qualitat del programa i de la seva eficàcia pedagògica. Al mateix temps dóna a conèixer a cada alumne el nivell de coneixements assolits, en relació als criteris establerts i a la resta del grup classe.

El marc de l'avaluació interna coincideix amb el marc de la tradicional avaluació que tots els professors han fet sempre dels alumnes i del seu propi treball.

Els projectes d'ensenyament dels últims anys han tingut especial cura en el disseny de l'avaluació que ha de realitzar el professor a l'aula. Els primers projectes que van presentar models d'avaluació sistemàticament estructurats per valorar tots els àmbits d'aprenentatge que proposaven van ser els models de renovació curricular dels anys 60. Entre aquests models destaquen nombrosos projectes d'ensenyament de les ciències: el Nuffield a Gran Bretanya, els projectes SAPA (Science A Process Approach), ESS (Elementary Science Study) i SCIS (Science Curriculum Improvement Study) a EE.UU. etc... La característica comuna a tots ells és l'interès per definir amb precisió els coneixements que l'alumne ha d'adquirir i la voluntat d'analitzar aquests coneixements en l'avaluació.

Un exemple de l'avaluació interna corresponent a aquesta època pot ser el disseny d'avaluació del programa SCIS.

El programa d'ensenyament elemental de les ciències SCIS va néixer als EE.UU. a la dècada de 1960 arran del moviment de reforma curricular. L'any 1976 es va fer una avaluació exhaustiva dels resultats de la seva experimentació i es va realitzar una remodelació profunda de les activitats, conceptes, temes,... que va conduir a la proposta d'un programa nou que s'anomenà SCIS. En aquest nou plantejament l'avaluació del rendiment dels alumnes es fonamenta en les observacions que fa el professor al llarg de la classe. Els aspectes que es proposen avaluar mitjançant l'observació són: les actituds, les habilitats i el treball de cada alumne. Per realitzar les observacions el professor disposa d'una guia de comportaments a observar (curiositat, pensament crític,...)

Per avaluar la comprensió de conceptes i de processos utilitzen proves de llapis i paper amb preguntes tancades o obertes, sempre acompanyades d'una estricta definició dels criteris de valoració. En tots els casos les preguntes fan referència a activitats desenvolupades a l'aula, i cada alumne (o bé en grups reduïts) les ha d'interpretar i justificar. Els alumnes que presenten dificultats en alguna prova passen a una entrevista individual amb el professor per resoldre les dificultats que han demostrat. També es proposa que

els alumnes valorin la classe de ciències, a partir de fitxes amb preguntes sobre la percepció que cada alumne té o d'una determinada sessió de treball. Aquest enfoc de l'avaluació correspon a un model d'ensenyament de fa més de deu anys, però ja planteja la manera d'avaluar el programa d'ensenyament.

Després de veure algunes de les propostes més significatives sobre l'avaluació de l'ensenyament de les ciències no podem acabar sense parlar d'una proposta per al futur; l'avaluació de les ciències que ens ofereix el projecte del Ministeri d'Educació i Ciència del nostre país per a l'ensenyament secundari:

Les orientacions per a l'avaluació de l'ensenyament de les ciències en el Disseny Curricular Base del M.E.C. veuen l'avaluació com una recollida d'informació sobre el progrés dels alumnes en l'àmbit conceptual, procedimental i actitudinal, però també consideren imprescindible l'avaluació de tots els factors que intervenen en el procés docent: l'actuació del professor, els materials utilitzats, les activitats realitzades, l'ambient de treball a l'aula i el nivell de relacions humanes.

Defineix les següents categories d'avaluació:

1. Comprensió i Expressió.
2. Utilització de fonts d'informació.
3. Ús d'instruments de laboratori i de camp.
4. Coneixement i aplicació de les idees bàsiques de la Ciència.
5. Treball en grup.
6. Actituds.

Les tècniques d'avaluació que aconsellen són:

- L'observació directa de les actituds, les habilitats, els hàbits de treball, la participació i el progrés conceptual.
- El seguiment del quadern de treball de cada alumne que permet controlar l'expressió escrita, l'ús de fonts d'informació, etc...
- Les proves de llapis i paper, que aconsella que siguin de resposta breu o tests objectius perquè els valors que es manifesten en les preguntes obertes poden avaluar-se amb l'observació.
- Totes les activitats realitzades a la classe.
- L'autoavaluació.

Les categories d'avaluació d'aquest projecte es poden classificar en l'aprenentatge de coneixements, d'habilitats i d'actituds i per tant coincideixen plenament amb les actuals tendències d'ensenyament de les ciències. Les tècniques d'avaluació es troben en la línia pragmàtica d'aplicar tots els mitjans

possibles per recollir informació sobre el rendiment dels alumnes. Queden pendents d'aclarir les tècniques que podran aplicar-se per a l'avaluació del procés docent.

L'avaluació en els propers 10 anys

L'avaluació de l'ensenyament de les ciències es troba en un moment de propostes teòriques interessants i suggestives. En nombrosos països es treballa en programes de recerca avaluativa amb resultats molt positius. Les dificultats plantejades per l'avaluació del treball pràctic, la valoració d'actituds, de processos, etc... es resolen amb creativitat i rigor científic.

Els treballs de recerca en el camp de l'avaluació de les ciències tenen una immediata aplicació a la pràctica docent, que està demanant amb insistència, metodologies d'avaluació adequades als actuals models d'ensenyament. Cal una major difusió dels treballs publicats en els altres països i la coordinació de la recerca que es realitza en el nostre país.

Perque els models experimentals es transformin en realitats escolars fa falta que els educadors disposin de tècniques d'avaluació senzilles, fàcils d'utilitzar a l'aula i que reunixin les qualitats que han de tenir tots els instruments d'avaluació:

- objectivitat: en un nivell suficient perquè la seva interpretació sigui independent de la persona que aplica la prova.
- validesa: perquè mesuri els aprenentatges que el professor vol mesurar.
- fiabilitat: és una qualitat intrínseca de cada instrument d'avaluació que garanteix l'obtenció de resultats estables.

Els models d'avaluació quantitativs, tan partidaris dels valors numèrics han desenvolupat sistemes d'anàlisi estadística que permeten comprovar el nivell deficiència i validesa de les proves objectives; en canvi, els models qualitativs, més interessants en descriure els aprenentatges que en mesurar-los, no s'han preocupat massa de demostrar la seva capacitat com a instruments d'avaluació. En el moment present es fa palesa la necessitat de comprovar que la informació recollida amb instruments qualitativs és objectiva, fiable i vàlida. D'aquesta necessitat han sorgit noves maneres de demostrar la qualitat d'un instrument d'avaluació: la triangulació, la contrastació de dades entre diferents avaluadors i diferents instruments, etc...

L'elaboració d'instruments d'avaluació dels conceptes, tècniques i actituds dels alumnes en l'àrea de ciències és una tasca imprescindible i urgent, que requereix un esforç important d'investigació i validació de les tècniques d'avaluació adequades als tres tipus d'aprenentatge. El professor ha de disposar

d'abundant material didàctic que li permeti de seleccionar el més convenient per al seu programa d'ensenyament, amb la garantia d'un nivell de qualitat comprovada. Si els professors no disposen d'instruments d'avaluació experimentals i pràctics serà difícil superar l'estadi de l'avaluació intuïtiva.

Les dificultats que sorgeixen en l'elaboració de proves de ciències són de dos tipus:

- Els problemes derivats de la complexitat dels coneixements que s'han d'avaluar: l'avaluació de processos independents dels continguts, l'avaluació de les habilitats d'investigació, l'avaluació d'actituds,...
- Els problemes tècnics que s'originen en la construcció i validació de les proves: la fiabilitat de les proves, la seva interpretació, el nombre de proves i la seva durada,...

Altres qüestions fan referència al paper de l'avaluació, per exemple la separació entre avaluació i ensenyament és, cada vegada, més difusa perquè l'avaluació individual dels alumnes dins del programa d'ensenyament té una funció formativa que no es pot diferenciar de l'educació de l'estudiant. També s'està estudiant la necessitat de controlar l'avaluació educativa per garantir la seva qualitat, per aquest control s'hauran d'idear sistemes de metaavaluació.

L'interès que la problemàtica avaluativa ha despertat en el professorat de ciències permet pensar que en els propers anys es produiran importants progressos en aquest camp. L'avenç en l'avaluació assegura l'avenç en l'ensenyament, perquè cal tenir en compte que un concepte nou d'ensenyament de les ciències no pot arrelar si l'avaluació segueix el model anterior. La importància de l'avaluació va més enllà del valor autotèlic perquè és un factor determinant de la possibilitat d'innovació en l'ensenyament.

BIBLIOGRAFÍA

- ALBERTS, R.V.J.; van BEUZEKEN, P.J.; de ROO, I. «The assessment of practical work: A choice of options» a *European Journal of Science* vol. 8, pp 361-369. 1986.
- ALVÁREZ, J.M. «Evaluando la Evaluación» a *Didáctica, currículo y evaluación: Ensayos sobre cuestiones didácticas*. Alamex. Barcelona, 1985.
- APU. *Science at age 13 and 15. Sample questions*. Department of Education and Science. London, 1985.
- APU. *Design and Technological activity*. Department of Education and Science. London, 1987.

- AUSUBEL, D.P. *Psicología educativa. Un punto de vista cognoscitivo*. Trillas. México, 1978.
- BERNARDO CARRASCO, J. *Cómo evaluar el aprendizaje*. Anaya. Madrid, 1985.
- BLOOM, B.S. i col. *Evaluación del aprendizaje. Ciencias*. Tomo 3. Troquel. Buenos Aires, 1975.
- BRYCE, T.G.K. et al. *TAPS 2. Techniques for the assessment of practical skills in Foundation Science*. Heinemann. London, 1987.
- BRYCE, T.G.K.; ROBERTSON «What can they do? A review of practical Assessment in Science» a *Studies in Science Education*, 12, pp 1-24. 1985.
- COLL, C. *Psicología y curriculum*. Laia. Barcelona, 1987.
- COMBER, L.C. i KEEVES, J.P. *Science Education in nineteen countries*. Stockholm, Almqvist and Wiksell. John Wiley. London, 1973.
- COOK i REIHART. *Métodos cuantitativos y cualitativos de investigación evaluativa*. Morata. Madrid, 1986.
- CRONBACH, L. *Toward Reform of Program Evaluation*. Jossey-Bass. San Francisco, 1980.
- DELTOUR, N. *Sciences: Dossier Technique*. Université de Liege. Liege, 1983.
- DRIVER, R.; CHILD, D.; GOTT, R.; HEAD, J.; JOHNSON, S.; WORSLEY, C.; WYLIE, F. *Science in schools. Age 15: Raport no. 2*. HMSO for the APU. London, 1984.
- DRIVER, R.; GUESNE, E.; TIBERGHIE, A. *Ideas científicas en la infancia i la adolescencia*. Morata/MEC. Madrid, 1989.
- EISNER, E. *The methodologie of Qualitative Evaluation*. Stanford University, Unpublish. 1981.
- ELLIOT, J. *Self Evaluation Professional Development and Account ability*. Cambridge Institute of Education. 1982.
- FRASER, B.J. «Development of a test of Science-related attitudes» a *Science Education*, 62 (4) pp 509-515. 1978.
- GALAN, M.I.; MARIN, D.E. *Investigación para evaluar elcurriculo universitario*. Universidad Nacional Autónoma de México. México D.F., 1988.
- G.C.S.E. *Your guide to the new exams*. Scottish Standard Grade (14-16). Glasgow. 1986.
- G.C.S.E. *Draft Syllabus in Human Biology for the 1988 Examination*. Southern Examining Group. Bristol, 1985.
- G.C.S.E. *Draft Syllabus in Applied Science for the 1988 Examination*. Southern Examining Group. Bristol, 1985.
- G.C.S.E. *Biology*. Northern Examining Association. York, 1988.
- G.C.S.E. *Chemistry data book*. Northern Examining Association. York, 1988.

- G.C.S.E. *Physics, 1989 Examination*. University of Oxford. Delegacy of local Examinations. 1987.
- G.C.S.E. *Nuffield co-ordinated Sciences. Examination Byllabuses 1990*. Mioland Examining Group. Cambridge, 1988.
- GIL, D.; GONZÁLEZ, A.; AGUILAR, R. «L'ensenyament de la física a través de les proves d'avaluació» a *Begones jornades de recerca educativa*. Lleida, 1986.
- GIORDAN, A. «Interés didáctico de los errores de los alumnos» a *Enseñanza de las ciencias*, 3(1) pp 11-17. 1985.
- GOTT, R.; MURPHY, P. *Assessing investigations*. Science report for teachers: 9. Assessment of Performance Unit. London, 1987.
- GUBA, E.; LINCOLN. *Effective Evaluation*. Jossey-Bass. London, 1985.
- HARLEN, W. *Guides to Assessment in Educator: Science*. Macmillan. London, 1983.
- HARLEN, W. *Enseñanza y aprendizaje de las ciencias*. Moratas. Madrid, 1988.
- HEWSON, P.N. «A case study of Conceptual Change in Special Relativity: The influence of Prior Knowledge in Learning» a *European Journal of Science Education*, 4(1), pp 61-78. 1982.
- HILL, L. «Biology, Philosophy and Scientific method» a *Journal of Biological Education*, 19 (3), pp 227-231. 1985.
- JOHNSON, S. Assessment in Science and Technology a *Studies in Science Education*, 14, pp 38-108. 1987.
- KEMPA, R. *Assessment in Science*. Cambridge University Press, 1986.
- LAND, H.G. «Criterion-referenced tests in Science: an investigation of reliability, validity and standards-settint» a *Journal of Research in Science Teaching*, 19(8), pp 665. 1982.
- MacDONALD, B. «La evaluación y el control de la educación» a *La enseñanza: su teoría y su práctica* de Sacristán/Pérez. Akal. Madrid, 1983.
- MAGER, R.F. *Medición del intento educativo*. Guadalupe. Buenos Aires, 1975.
- M.E.C. *Diseño curricular base. Educación primaria*. Ministerio de Educación y Ciencia. Madrid, 1989.
- M.E.C. *Diseño curricular base. Educación Secundaria* Ministerio de Educación y Ciencia. Madrid, 1989.
- METFESSEL, N.S.; MICHAEL, W.B. «A paradigm involving multiple criterion measures for the evaluation of the effectiveness of school programs» a *Educational Psychological Measurement*, 27, pp 931-943. 1967.
- NAEP. *Selected results from the National Assessment of Science: scientific principles and procedures*. Report no 04-5-02. Education of the States. Washington, 1975.

- NAEP. *Attitudes towards science*. Report no 08-8-02. Education Commission of the States. Washington, 1979.
- NAEP. *Three assessments of Science, 1969-1977: Technical summary*. Education Commission of the States. Washington, 1979.
- PARLETT, M.; HAMILTON, D. «Illuminative Evaluation» a *Human Inquiry: A source of new paradigm research*. Wiley. London, 1984.
- PÉREZ GÓMEZ, A. «Modelos contemporáneos de evaluación a *La enseñanza: su teoría y su práctica* de Gimeno y Pérez. Akal. Madrid, 1983.
- PIDGEON, D.; YATES, A. *Evaluación y medida del rendimiento escolar*. Anaya. Madrid, 1979.
- POPHAM, W.J. *Evaluación basada en criterios*. Magisterio Español. Madrid, 1983.
- REGENTS OF THE UNIVERSITY OF CALIFORNIA. *S.C.I.I.S. Teacher's handbook*. University of California. Berkeley, 1976.
- SATTERLY, D. «Los exámenes referidos al criterio y al concepto en ciencias: un nuevo sistema de evaluación» a *Enseñanza de las ciencias* 6(3), pp 278-284. 1988.
- SCRIVEN, M. «The methodology of evaluation» a *Perspectives on curriculum evaluation*. Rand McNally. Chicago, 1967.
- SCRIVEN, M. «Pros and cons about goal-free evaluation» a *Evaluation Comment*, 3, pp 1-4. 1974.
- STAKE, R.E. «A theoretical statement of responsive evaluation» a *Studies in Educational Evaluation*, 2, pp 19-22. 1976.
- STENHOUSE, L. *La investigación como base de la enseñanza*. Morata. Madrid, 1987.
- STENHOUSE, L. *Investigación y desarrollo del currículo*. Morata. Madrid, 1987.
- STUFFLEBEAM, D.L.; SHINKFIELD, A.J. *Evaluación sistemática: Guía teórica y práctica*. Paidós/MEC. Madrid, 1987.
- WELFORD, G.; HARLEN, W.; SCHOFIELD, B. *Practical testing at ages 11, 13 and 15*. Science report for teachers: 6 for APU. London, 1985.
- YAGER, R.E. «Asses all five domains of Science» a *The Science Teacher*, octubre, pp 33. 1987.