

## Capítol 4. L'avaluació de les competències

---

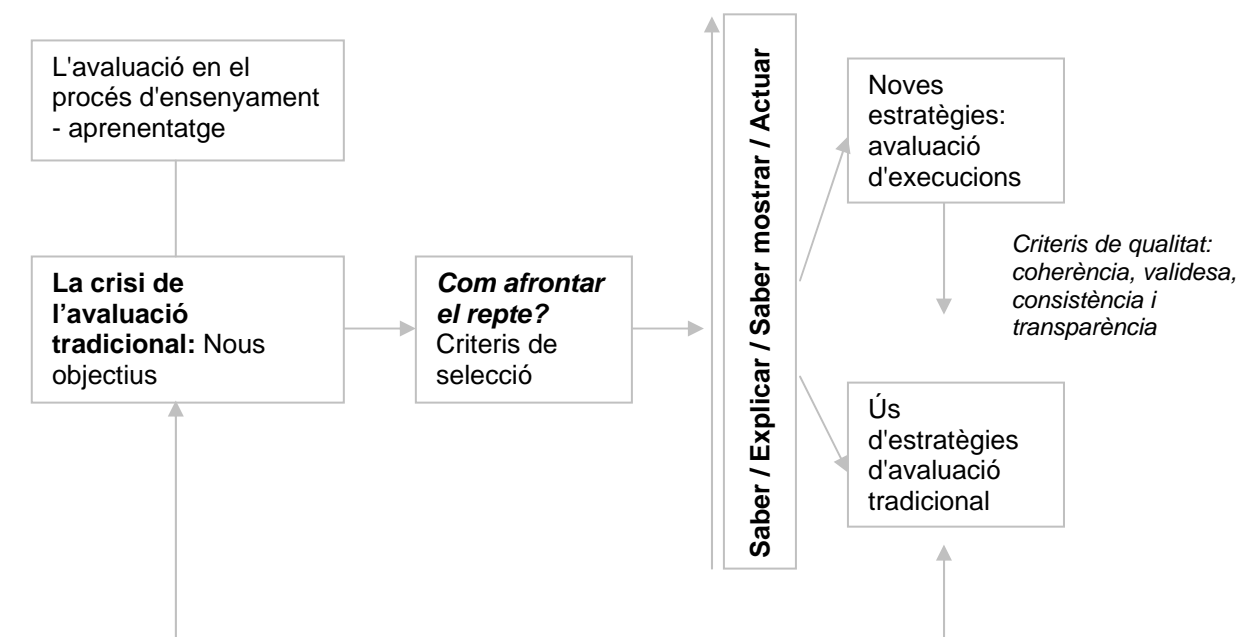
Capítol 4. L'avaluació de les competències .....	105
4.1. L'avaluació: una fase en el procés d'ensenyament - aprenentatge .....	107
4.2. La crisi dels sistemes tradicionals d'avaluació i noves tendències .....	111
Nous objectius, noves formes d'ensenyar i d'aprendre, noves formes d'avaluar.....	111
Els procediments tradicionals <i>versus</i> l'avaluació d'execucions.....	113
4.3. Criteris per seleccionar les estratègies d'avaluació .....	116
a) La coherència .....	116
b) La validesa i fiabilitat.....	118
c) Altres criteris: l'impacte en l'aprenentatge i el cost.....	120
4.4. Estratègies d'avaluació.....	122
a) L'avaluació tradicional.....	122
b) L'avaluació d'execucions .....	125
c) Conclusions.....	131
4.5. La gestió de l'avaluació .....	134

1	NOUS REPTES DE FORMACIÓ SUPERIOR
2	LA DEFINICIÓ DEL PERFIL DE FORMACIÓ I LES COMPETÈNCIES
3	APRENTATGE I FORMACIÓ DE COMPETÈNCIES
4	<b>AVALUACIÓ DE COMPETÈNCIES</b>
5	LES COMPETÈNCIES DEL PERFIL DE FORMACIÓ DE QUÍMICA

En els capítols previs s'ha argumentat com els canvis en l'entorn han introduït canvis en el currículum, canvis en els objectius d'aprenentatge a assolir, als quals es fa referència amb el terme de "competències". El canvi conceptual dels objectius d'aprenentatge implica un canvi en les estratègies docents, i en la concepció dels rols de professor i aprenent (capítol 3).

En aquest capítol es prosseguirà amb la cadena de canvis que genera la situació descrita al primer capítol. La introducció de nous objectius d'aprenentatge crea la necessitat d'ampliar o diversificar les estratègies avaluatives, de manera que es pugui obtenir informació sobre l'assoliment en aquests nous resultats d'aprenentatge.

L'objectiu, doncs, d'aquest capítol és el de presentar enfocaments avaluatius pertinents per avaluar les competències. La lògica que subjeu a l'estructura del capítol s'articula de la manera següent:



Esquema 4.1

## 4.1. L'avaluació: una fase en el procés d'ensenyament - aprenentatge

---

És àmpliament admès que l'avaluació és un procés sistemàtic de recollida d'informació (mesurar), emissió d'un judici de valor (comparar) i presa de decisions (certificar o no certificar). En aquesta tesi se sosté, a tall d'hipòtesi, que l'anàlisi dels sistemes avaluatius és una aproximació adequada per al coneixement de les competències que es desenvolupen en un determinat ensenyament<sup>1</sup>, i això és perquè l'avaluació no és el darrer estadi en un procés d'ensenyament, sinó que té importància des del primer moment.

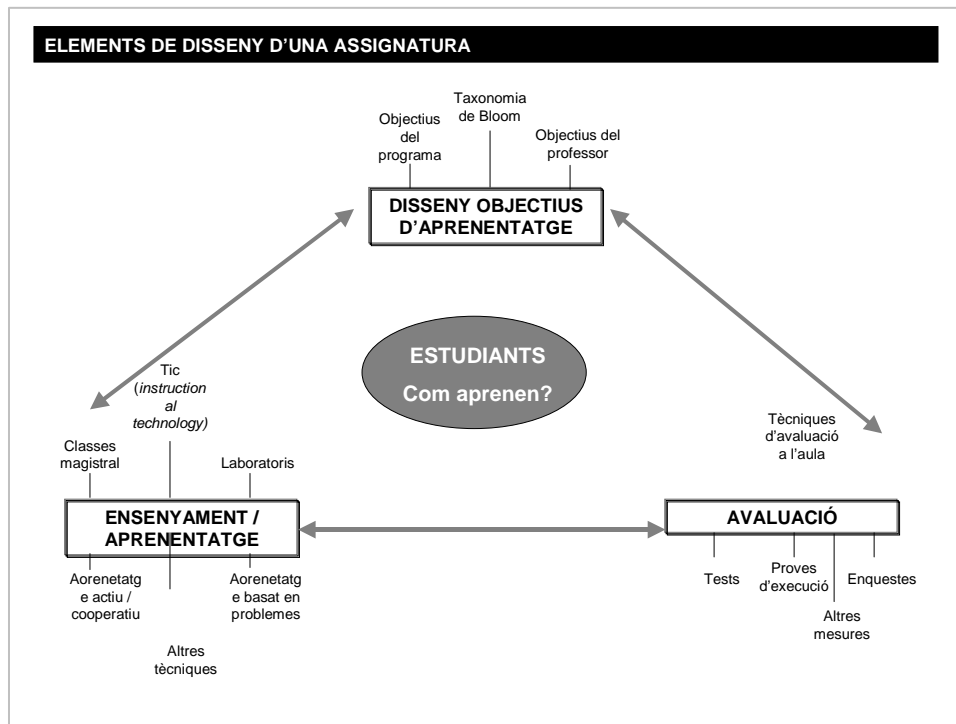
La fase avaluativa s'alimenta de la fase d'ensenyament aprenentatge: així, les metodologies d'avaluació s'han de dissenyar per tal que puguin mesurar l'assoliment dels objectius; a la vegada, el procés avaluatiu alimenta el procés d'ensenyament aprenentatge, i dóna *feedback* a estudiants i professors sobre l'assoliment d'objectius, identificació de llacunes, concepcions errònies, etc.<sup>2</sup>

L'avaluació ha de permetre determinar punts de referència que assenyalin la pertinència o no que l'estudiant pugui avançar en la complexitat que li suposa tot programa de formació (presa de decisions), així com certificar l'assoliment dels estudiants en els objectius o estàndards pretesos. A més a més, un pla d'avaluació hauria de comportar i guiar la reflexió sobre el procés d'ensenyament aprenentatge, implicant tant a professorat com a alumnat (Brown i Knight, 1994).

---

<sup>1</sup> Samuelowicz i Bain ([1992] citats a Bennet, Dune i Carré, 2000) diferencien entre la concepció *ideal* de l'ensenyament i la concepció *de treball* de l'ensenyament (*ideal vs working conception of teaching*). Aquesta diferenciació permetria explicar per què es produeix incoherència entre allò que expressen els professors, i allò que fan. Així, per exemple, els professors poden manifestar el seu desacord en el fet que els principals propòsits de l'avaluació són donar notes sumatives, però en canvi emprar procediments tradicionals d'avaluació on la preocupació real són les notes sumatives (Bennet, Dune i Carré, 2000). Sembla, doncs, que si cal triar entre analitzar com avaluen o com diuen que avaluen, és més adequat analitzar els procediments avaluatius.

<sup>2</sup> De fet, els tres tipus d'avaluacions –diagnòstica, formativa i sumativa– tenen aquesta funció de retroalimentar el procés d'ensenyament aprenentatge, tot i que potser no en el mateix grau. Així, l'avaluació diagnòstica dóna informació al professorat sobre el nivell de coneixements previs dels estudiants, informació que ha de servir per dissenyar el procés formatiu prenent en consideració aquesta situació. Per la seva banda, l'avaluació formativa no té altre objectiu que el de millorar el procés d'ensenyament aprenentatge mitjançant donar informació tant a professorat com a estudiants sobre el progrés dels aprenentatges. Finalment, l'avaluació sumativa permet valorar l'efectivitat del procés d'ensenyament i, tot i que ja no s'està a temps de modificar el procés d'aprenentatge, permet reorientar els processos futurs (Espín, Rodríguez, 1993).



Esquema 4.2. Font: adaptat de Felder, R; Brent, R (2003)

L'esquema 4.2 il·lustra aquesta interrelació de les diferents fases - perfil o objectius d'aprenentatge, metodologies docents i avaluació –i com els tres processos han de pivotar sobre el coneixement de com aprenen els estudiants, descrit amb més detall al Capítol 3 (motivació, coneixements previs, etc.). La forma d'avaluar pot determinar la manera en la qual els estudiants **enfronten** el seu aprenentatge.

De fet, mentre que el professorat sol concebre l'avaluació com el darrer pas en el procés ensenyament aprenentatge, per a l'estudiant és el primer element, l'element a partir del qual organitzarà les seves estratègies d'aprenentatge (CSHE, 2002). L'avaluació té un paper fonamental per a fixar els continguts i els nivells (tant d'esforç com de processament cognitiu: memorització, comprensió, etc.) que seran objecte d'atenció prioritària, que condicionen de forma molt determinant els estils d'aprenentatge dels estudiants<sup>3</sup>.

<sup>3</sup> Així, per exemple, encara que un professoramenti que és molt important l'avaluació de mètodes oposats o alternatius, però els seus exàmens consisteixen en la descripció i no en l'avaluació crítica dels mètodes, l'estudiant no desenvoluparà les habilitats d'avaluació crítica (Miller, Imrie i Cox, 1998).

No és tan sols un element dinamitzador de la qualitat dels aprenentatges (AQU, 2003), també ho és de l'activitat docent. L'avaluació té repercussions en el procés i el producte, no només en la fase final (Espín i Rodríguez, 1993; Havnes, 2004): alumnes i professors modifiquen la seva conducta en funció d'allò demanat per l'avaluació, només cal pensar en l'educació de l'antic COU, clarament orientada a les PAAU, o la influència que exerceixen els exàmens MIR en l'ensenyament de Medicina<sup>4</sup>.

Per tant, una part considerable de l'èxit del procés d'ensenyament - aprenentatge depèn del bon disseny de l'avaluació, de la seva coherència amb els objectius d'aprenentatge, de dissenyar l'avaluació "com si l'aprenentatge fos el més important"<sup>5</sup>. Perquè sigui coherent cal que les demandes avaluatives estiguin connectades amb l'organització i desenvolupament del currículum (QAA, 2000a).

### **La preocupació dels "organismes"**

El canvi de paradigma, d'avaluar inputs o processos a avaluar resultats, es tradueix en el fet que els organismes competents en l'avaluació o acreditació d'institucions d'ensenyament superior incrementin la preocupació sobre quins són els mecanismes d'avaluació dels ensenyaments i quin ús es fa dels resultats. L'anàlisi dels resultats d'avaluació té, doncs, una doble vessant: com a mitjà de retiment de comptes mitjançant la transparència dels resultats, però com a element indispensable per a la gestió de l'ensenyament, ja que l'anàlisi d'aquests resultats és un element clau per a la diagnosi de punts forts i febles, i l'elaboració consegüent d'un pla de millores.

Així, per la Comission on Higher Education (1997) l'avaluació és l'element clau per millorar l'efectivitat de l'ensenyament, ja que permet guiar la reflexió sobre el procés d'ensenyament d'aprenentatge, assegurar la millora institucional, i, finalment, facilitar el retiment de comptes.

---

<sup>4</sup> Un altre cas paradigmàtic dels efectes indirectes de l'avaluació el relata Anton Havnes (2004) que analitza, mitjançant un estudi etnogràfic, un curs semestral obligatori per a tots els alumnes que ingressen a la Universitat d'Oslo que té com a objectiu fomentar una aproximació independent i crítica de la ciència, però que acaba sent una preparació per als exàmens finals. Així, els estudiants es troben amb unes declaracions d'objectius que no es veuen reflectides ni en la tipologia de classes (metodologies docents), ni en els manuals, ja que tant les unes com els altres estan orientats a la superació dels exàmens.

<sup>5</sup> Thomas A Angelo, en un article, significativament titulat *Doing Assessment As If Learning Matters Most* (1999), assenyala que la funció més important de l'avaluació és fomentar l'aprenentatge; i, per tant, l'avaluació ha de fomentar les habilitats, disposicions i coneixement necessaris per a implicar-se activament (intel·lectualment i emocional) en el treball acadèmic, esdevenir conscient de les preconcepcions, creences i coneixements previs i estar disposat a desaprendre quan sigui necessari, o comprendre i valorar els criteris, estàndards i mètodes pels quals són avaluats.

Per la seva banda, la britànica Quality Assurance Agency (2000a) estableix que l'avaluació té diverses funcions: és el fonament de la presa de decisions de si l'estudiant està preparat per progressar, permet que els estudiants obtinguin feedback del seu aprenentatge ajudant-los a millorar-lo, i possibilita que el professorat avalui l'efectivitat del seu ensenyament.

L'Australiana CSHE (Center for the Study for Higher Education, 2002), precisa que, per tal que l'avaluació pugui ser una eina estratègica per a la millora del procés d'ensenyament aprenentatge, cal que l'avaluació fomenti aproximacions efectives de l'aprenentatge i mesuri de manera vàlida i fiable els resultats esperats (particularment els d'alt ordre que caracteritzen l'educació superior).

En el nostre context també es recullen com a funcions de l'avaluació aquests tres potencials: la millora de l'aprenentatge dels estudiants, la comprovació de l'assoliment d'objectius pel que fa a la superació d'un estàndards i la d'optimització de la docència, proporcionant informació que l'avaluació aporta per a la gestió de l'ensenyament (AQU, 2003a).

## 4.2. La crisi dels sistemes tradicionals d'avaluació i noves tendències

---

### **Nous objectius, noves formes d'ensenyar i d'aprendre, noves formes d'avaluar**

L'excel·lència universitària rau en la qualitat i quantitat dels nivells d'assoliment dels aprenentatges dels estudiants, i per a mesurar-ne i valorar-ne l'excel·lència és clau disposar de sistemes de recollida d'informació, és a dir, de l'adequació dels procediments avaluatius. Per tant, si es canvia l'objectiu de formació, ja sigui pel que fa a la tipologia d'allò a aprendre o a la quantitat, cal canviar els sistemes d'avaluació capaços de copsar l'assoliment dels nous objectius. Caldrà dissenyar noves metodologies que capturin informació sobre el rendiment de l'aprenentatge, si es pretén capturar informació substantiva als nous objectius (Mateo, 2001).

Si s'assumeix el concepte de competències descrit al Capítol 2, no només han de canviar les pràctiques educatives (Capítol 3), també ho hauran de fer les avaluatives. A continuació s'assenyalen algunes implicacions:

- Cal que es produeixi una aprenentatge profund, s'hauran d'evitar aquells instrument d'avaluació que afavoreixen les estratègies superficials d'aprenentatge, és a dir, els que fomenten el simple record (Simpson, 2003).
- Cal assegurar que a les metodologies formatives els correspongui un pes avaluatiu<sup>6</sup>.
- El constructe de competències és un constructe integrador, de coneixements habilitats i actituds, per tant, les proves avaluatives hauran d'oferir oportunitats per a exhibir aquesta integració.
- Les competències es demostren, s'avaluen en execucions o demostracions, per tant, les avaluacions d'execucions prenen importància

---

<sup>6</sup> En un model docent, com el que implica la introducció dels crèdits europeus (ECTS), en el qual l'avaluació s'ha de dur a terme en base al treball de l'estudiant, no sembla coherent que hi hagi demandes, per exemple, de treball en grup però que aquestes no s'avaluin o que en el seu pes en l'avaluació final no hi hagi certa proporcionalitat en relació a la quantitat de temps que ha implicat.

- Les competències són contextuals, per tant, a més de l'adquisició de les competències (coneixements, habilitats disciplinars, competències transversals), cal avaluar el coneixement de quan i com aplicar-los (Dietel, Herman, Knuth, 1991).
- La competència és un constructe dinàmic, no és una escala de tot o res, sinó que indica l'evolució del novell a l'expert. Avaluar una competència ha d'implicar avaluar-ne el desenvolupament.
- Un dels nuclis importants de la competència és l'autonomia. Caldrà emprar eines que afavoreixin la progressiva autonomia, que fomentin l'autodiagnòstic realista i la recerca proactiva de recursos per a assolir els objectius establerts. L'aprenent s'ha de responsabilitzar o corresponsabilitzar del seu propi aprenentatge, ja que, en definitiva, l'important és que aprengui, la qual cosa és absolutament individual. L'autoreflexió és la base de l'aprenentatge al llarg de la vida, un dels principals desafiaments que afronta l'educació<sup>7</sup>.

Avaluar resultats d'aprenentatge, entesos com a competències, implica, per tant, diversificar les tecnologies o els instruments de recollida d'informació, i, a més a més, diversificar els agents d'avaluació. Així, es promou que els propis estudiants avaluin el seu aprenentatge, mitjançant eines com els dossiers d'aprenentatge, que avaluin els seus companys, o fins i tot que els ocupadors o tutors de pràctiques esdevinguin avaluadors. La complexitat de les demandes avaluatives, els problemes de fiabilitat que algunes d'aquestes noves tecnologies impliquen, obliguen que l'avaluació no pugui ser una activitat individual del professor, i que, sovint, impliquin juntes d'avaluació.

Així doncs, és necessari que hi hagi **coherència** entre els objectius perseguits per la metodologia d'ensenyament i les formes d'avaluar l'assoliment dels objectius, per tal d'assegurar la funció del retiment de comptes<sup>8</sup>.

---

<sup>7</sup> La Confederació d'Associacions d'Ocupadors Alemanys (BDA, 2003) assenyala l'habilitat de reflexionar sobre les pròpies accions i el comportament necessari per a desenvolupar el ple potencial professional.

<sup>8</sup> És a dir, fent el símil amb els processos d'avaluació institucional, d'una banda cal assegurar la rellevància dels objectius de formació (*fitness of purpose*), cosa que, com s'ha vist, es el que es pretén amb aquest canvi d'enfocament, i de l'altra que aquests objectius realment s'assoleixen (*fitness for purpose*) (Gola, 2003). Per tant, cal ser curosos tant en l'establiment d'objectius rellevants i mesurables, com en l'elecció dels mètodes d'avaluació, que hauran de ser vàlids en relació a aquests objectius.



## Els procediments tradicionals *versus* l'avaluació d'execucions

Els procediments tradicionals d'avaluació entren en crisi: l'avaluació basada en assoliments (**outcome-based evaluation**) reclama, com a mínim, en una part significativa, una **performance-based assessment** (avaluació d'execucions). Aquesta avaluació requereix que els estudiants, en un determinat moment, demostrin, construeixin o desenvolupin un producte o solució a partir d'unes condicions definides i d'uns estàndards (Rodríguez Espinar, 2000). Les avaluacions basades en execucions en contextos aproximats als que els estudiants hauran de trobar-se en les seves vides personals i professionals, requereixen no només coneixements factuais, també habilitats, requereixen la integració de coneixement i habilitats en execucions complexes. Aquest tipus d'avaluació representa un canvi enorme en relació a les avaluacions estandarditzades, administrades a grans grups i on els estudiants sovint percebién poca relació entre allò que se'ls avalua i el que se'ls requeriria un cop estiguessin graduats (Wright, B;1997).

A continuació es resumeixen les noves tendències en l'avaluació<sup>9</sup>:

### a) Canvis en l'abast i els objectius de l'avaluació

- Es passa d'una avaluació centrada en el control a una de centrada en l'aprenentatge: avaluar ja no és només demostrar o certificar, és reflexionar, perfeccionar, optimitzar.
- Es passa d'avaluar habilitats discretes i isolades, a integrades i interdisciplinàries, és a dir, a avaluar "competències" i no només parts discretes d'aquestes.
- S'amplien els àmbits de l'avaluació, no només s'avaluen **assoliments acadèmics i professionals**, sinó també **assoliments personals** (competències genèriques i transferibles).
- Es passa d'avaluar l'acumulació del coneixement, a avaluar-ne l'aplicació, l'ús dels fets i habilitats apreses, desplaçant l'èmfasi de la reproducció al pensament d'alt ordre (anàlisi i aplicació, síntesi, avaluació).

---

<sup>9</sup> Per elaborar aquest apartat s'han emprat els articles/líbrs següents, les cites dels quals s'han eliminat del text per tal d'alleugerir-ne la lectura: Bordas i Cabrera, 2001; Dietel, Herman i Knuth, 1991; Heywood, 2000; Mateo, 2001; Rodríguez Espinar, 2000; Wright, 1997.

- S'incrementa, per tant, la complexitat i amplitud de les tasques d'avaluació, de manera que es pugui assegurar que les tasques d'avaluació siguin representatives del perfil de formació o dels objectius d'aprenentatge pretesos. Des d'aquesta perspectiva, no té gaire sentit emprar molts exàmens escrits que avaluen la mateixa competència, sinó que l'avaluació s'ha de fer des d'estratègies múltiples.

## b) Canvis en les característiques de les proves d'avaluació:

- Pas de les **proves de paper i llapis, a avaluacions d'execucions**, on es plantegen problemes contextualitzats la resposta dels quals és rellevant i significativa. De l'extensivitat de les proves objectives d'elecció múltiple a proves que reclamin la construcció de la resposta. Per tant, les formes d'avaluació cada cop empen més aproximacions qualitatives, a banda de les quantitatives.
- Es diversifiquen les fonts d'evidència sobre els judicis en relació a la qualitat de l'aprenentatge de l'estudiant. No només es pren en compte la informació factual o de conceptes, sinó que també s'avaluen de manera multidimensional els nivells de destreses, i d'altres qualitats com ara la creativitat, la persistència, la meticulositat, l'empatia, etc.
- Pren importància l'avaluació **formativa** que es realitza de manera integrada al propi procés d'ensenyament aprenentatge que s'avalua<sup>10</sup>. L'avaluació formativa es basa en els productes o mostres de treball pròpies del desenvolupament normal d'un curs; s'empen, no només per assignar una puntuació, sinó per avaluar què reflecteixen en relació a l'eficiència del procés d'ensenyament aprenentatge, la qual cosa, a més de

---

<sup>10</sup> De fet les avaluacions titllades "d'objectives" sovint han demostrat tenir molt poca validesa: W.D Furneaux (1962, citat a Heywood, 2000) va analitzar els resultats dels exàmens duts a terme en un ensenyament d'enginyeria tècnica. La hipòtesi era que els exàmens de primer eren un requisit per prosseguir segon, de manera que la nota dels exàmens de primer estaria correlacionada amb la dels de segon. Tanmateix, es va trobar molt poca relació entre les assignatures (cap era superior a .35), i no sempre l'esperada (la correlació de matemàtiques de 2n amb les de 1r era més baixa que la de matemàtiques de 2n amb electricitat aplicada a 1r). D'altra banda, l'anàlisi factorial dels resultats dels exàmens de primer van mostrar un únic gran factor, de la qual cosa es desprèn que cadascun dels exàmens de primer mesurava un tipus d'habilitat comuna per superar exàmens. A partir d'aquest estudi, es van substituir els exàmens de darrer any de carrera per diferents mòduls: un de coneixement i principis (examen tipus test), un de comprensió d'enginyeria a través de preguntes d'opinió sobre un article, un altre mòdul de disseny i planificació (replacar la planificació d'un projecte), i un mòdul de respostes llargues i curtes sobre l'aplicació dels principis de l'enginyeria amb problemes pràctics, alguns d'ells relacionats amb el context economicosocial. L'anàlisi factorial va mostrar un factor per cada subtest, cosa que confirma que, efectivament, s'estaven mesurant competències o dominis, o objectius d'aprenentatge diferents (Heywood, 2000:33-34)

permetre estalviar-se temps, fomenta la motivació dels estudiants en el seguiment del curs. Així, s'incorporen els projectes de treball (models de realitzacions) com objectes d'avaluació.

- D'altra banda, una de les funcions de l'avaluació que recobra importància és l'avaluació entesa **com a feedback per l'estudiant**. En situar l'èmfasi sobre el que aprenem, més que sobre el que s'ensenya, el rol de l'aprenent pren poder, cobra importància. Una implicació d'aquesta afirmació és la corresponsabilització dels alumnes en els processos avaluatius. Ser equitatiu, fer avaluacions fiables, no és tan important com fomentar l'autoconsciència de l'aprenentatge.

**c) Canvis en "qui" s'avalua:** de l'avaluació individual, a l'avaluació del procés grupal, de les competències interpersonals, o "habilitats col·laboratives".

**d) Canvis de qui és l'avaluador:** s'incrementen els agents avaluadors. A més del professor o d'una junta de professors, el propi estudiant, o els seus companys de classe, o bé a nivell extern de la institució els ocupadors, o *assessment centers*, poden esdevenir avaluadors. Inclusió de l'avaluació compartida.

**e) Canvis en la freqüència d'avaluació:** d'única, a obtenir mostres al llarg del temps (dossiers d'aprenentatge) que donen les bases per a l'avaluació. La superació de la concepció avaluació puntual, ja no és un acte final, sinó un procés paral·lel.

### 4.3. Criteris per seleccionar les estratègies d'avaluació

---

L'avaluació és una fase cabdal en el procés d'ensenyament aprenentatge. Atès que hi ha nous objectius d'avaluació, el criteri de coherència fa que entrin en crisi les estratègies d'avaluació tradicional. La crisi no és de substitució d'objectius, sinó d'increment de la seva complexitat. Les estratègies tradicionals no queden, per tant, apartades, però se n'afegeixen d'altres; s'hi troben, però, les tendències següents, que són coherents amb la introducció del constructe de competències:

Proves de paper	↔	Altre tipus de proves
Avaluació referida a la norma	↔	Avaluació criterial
Emmagatzament d'informació	↔	Comprensió i aplicació del coneixement

Davant de tanta diversitat avaluativa, una primera qüestió que apareix és com seleccionar els mètodes d'avaluació més pertinents als objectius desitjats. Els criteris més rellevants són la coherència amb els objectius d'aprenentatge, qüestions sobre la seva validesa i fiabilitat<sup>11</sup>, seguits d'altres qüestions relatives a l'impacte en l'aprenentatge i al cost en termes de temps i recursos (Miller, Imrie i Cox, 1998; AQU,2005).

#### a) La coherència

La introducció de nous objectius d'aprenentatge relacionats amb el saber fer, amb l'acció reflexiva, fa necessari introduir instruments d'avaluació que permetin avaluar-los. El criteri de coherència posa en entredit la suficiència de les estratègies d'avaluació tradicional per copsar informació sobre l'assoliment en les competències preteses<sup>12</sup>.

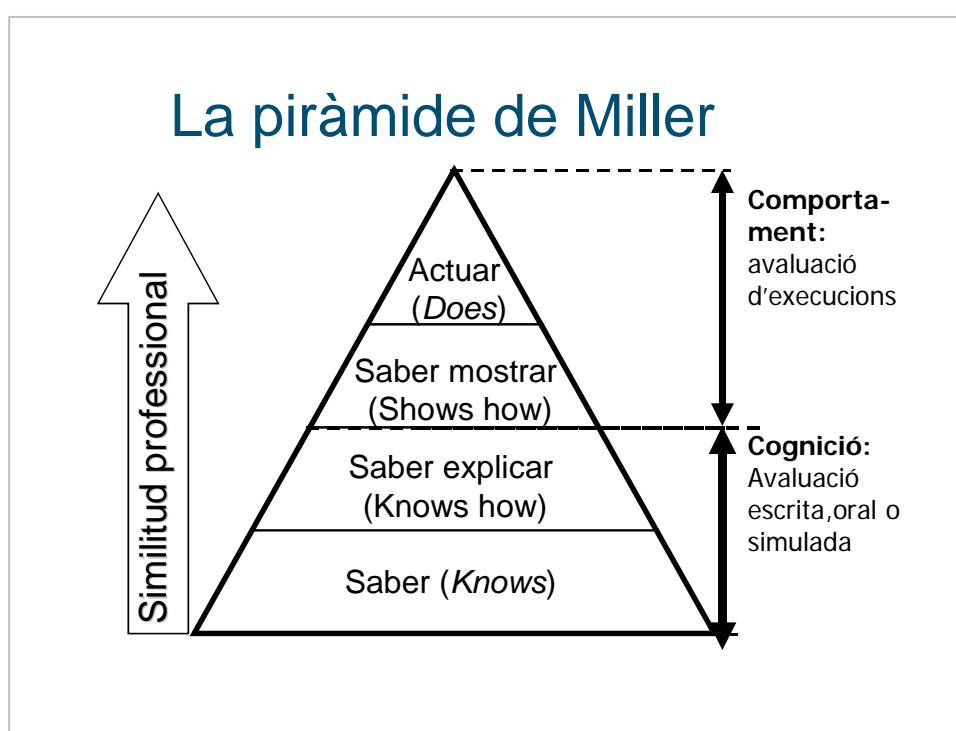
---

<sup>11</sup> Els estàndards de qualitat per a l'acreditació de les titulacions (AQU, 2005) assenyalen que tots el mètodes d'avaluació per determinar l'aprenentatge dels estudiants són vàlids, consistents i coherents amb els objectius d'aprenentatge establerts pel programa.

<sup>12</sup> La coherència està relacionada amb la validesa: estratègies d'avaluació coherents seran vàlides en relació a allò après. Però no només seran vàlides, també seran representatives. Així, es podria dissenyar una prova vàlida (per exemple, en una assignatura d'història, un assaig sobre el franquisme, però que no seria coherent si fos l'única prova sobre l'assignatura). Finalment, la coherència està relacionada amb el criteri d'impacte en l'aprenentatge: les estratègies d'avaluació coherents amb els processos d'ensenyament aprenentatge en reforcen la implicació (és a dir, la motivació en tot el procés).

D'altra banda, en els capítols 2 i 3 s'ha argumentat com les competències requereixen un enfocament d'aprenentatge profund (la qual cosa obliga a plantejar estratègies d'avaluació que, com a mínim, no el desanimin), així com que la competència es desenvolupa, la qual cosa vol dir que caldrà dissenyar avaluacions pels diferents estadis de desenvolupament.

La piràmide de Miller (1990) pot ser una manera útil d'ajudar a triar estratègies avaluatives coherents amb els objectius d'aprenentatge. Dins d'un perfil de formació, hi haurà continguts o elements que caldrà saber, d'altres que caldrà saber explicar, uns altres en els quals serà necessari saber com s'haurien de fer (saber simular), i finalment, hi haurà alguns elements que caldrà fer, dels quals –idealment– se'n certificarà l'execució (veure Figura 2).



Esquema 4.3. Font: Miller GE. (1990)

Seguint l'esquema 4.3, per les competències de nivell 1 i 2 (saber, i saber explicar; o des de la terminologia de Bloom, conèixer i comprendre) les estratègies d'avaluació tradicional són més que suficients (resulten vàlides), mentre que amb les de nivell 3 i 4 cal avaluació d'execucions o simulacions.

## b) La validesa i fiabilitat<sup>13</sup>

La **fiabilitat** es refereix a l'exactitud de la mesura, és a dir, a l'absència d'errors en la mesura. La fiabilitat fa referència a la consistència de les puntuacions obtingudes pels mateixos individus si fossin reexaminats amb la mateixa prova, en diferents ocasions, o amb proves amb qüestions equivalents, o amb condicions d'avaluació variables (Anastasi, 1976, citat a Miller; Imrie i Cox, 1998: 236).

La **validesa** fa referència al grau en què la mesura obtinguda reflecteix allò que es pretén mesurar<sup>14</sup>. La validesa d'un mètode d'avaluació està en funció del grau en què la prova mesura allò que pretén mesurar. Ara bé, per a poder-la establir cal haver definit clarament els objectius a assolir, cal que permeti una **avaluació criterial**.

Un dels errors més freqüents en l'elaboració de les proves sumatives consisteix a no establir una avaluació criterial en relació als objectius, la qual cosa fa que un estudiant aprovi un examen sense dominar un aspecte que serà necessari més endavant (fent la mitjana amb altres qüestions d'altres temes). Per cadascun dels blocs de contingut caldria determinar quins són aquells nuclears, és a dir, quins són aquells aspectes essencials i el grau en què caldrà dominar-los, ja sigui per a afrontar altres matèries, o bé per a la pràctica professional, i quins són complementaris o desitjables. Els primers haurien de ser criterials, és a dir, determinar l'aprovat vs el suspens, mentre que els segons són els que permetrien matisar la nota i valorar l'esforç afegit de l'estudiant.

Aquests dos criteris són els que permeten que no succeeixi el què adverteix Dressel:

La nota és un judici inadequat dut a terme per un jutge variable i esbiaixat sobre el grau en el qual un estudiant ha assolit un nivell mal definit de domini d'una proporció desconeguda d'un material indefinit.

(Dressel 1976, citat a Miller, Imrie i Cox 1998: 24)

---

<sup>13</sup> Aquest aspecte té un alt component psicomètric. Pot ser d'interès consultar els treballs següents que prenen com a objecte d'anàlisi els continguts de l'educació superior: Heywood (2000), Miller, Imrie i Cox, 1998; Palomba (1999)

<sup>14</sup> Howard Gardner (1993) assenyalava l'existència d'un tipus diferent de validesa, la **validesa ecològica**. Una avaluació en un entorn formal (per exemple, les proves de selectivitat) tenen menys validesa ecològica que la que es donaria en les situacions dels aprenents dels antics oficis, en les quals hi ha una formació supervisada, gradual; l'avaluació es realitza i modula al llarg del temps en el propi lloc de treball. Així doncs, no té gaire sentit qüestionar-se la validesa dels judicis del mestre de l'aprenent, ja que segur que podria predir amb bastant precisió les actuacions del seu aprenent.

## La fiabilitat i la validesa en les proves d'execució

Un dels problemes tècnics més greus que planteja l'avaluació de les proves d'execució és la seva fiabilitat, la qual cosa posa en perill la seva validesa. L'observació directa és una de les maneres més vàlides d'avaluar les habilitats tècniques o interpersonalment d'un estudiant que actua en un context real, simulat o d'examen (Cannon i Newble, 2000). Malauradament, la fiabilitat pot ser francament baixa, especialment en l'àrea interpersonal, on no hi ha cap altra forma d'avaluació. La fiabilitat de l'observació pot incrementar-se millorant l'instrument de puntuació a través de la construcció de *checklists* o d'escala de puntuació, o bé millorant l'execució de l'observador (Cannon i Newble, 2000).

Des de la investigació qualitativa s'ha treballat molt en la cerca d'estratègies que incrementin la fiabilitat (i per tant la validesa) de tècniques com ara l'observació o l'anàlisi de contingut, tècniques que es podrien traslladar, en l'àmbit de l'avaluació, en el procés d'ensenyament aprenentatge. A continuació se n'esbossen les principals<sup>15</sup>:

- Treball prolongat en el mateix lloc: permet tenir el coneixement de la situació des de les categories culturals dels qui hi viuen.
- Triangulació de temps o Observació persistent, triangulació de dades: diferents dades a partir de les quals s'emet un judici, triangulació d'estratègies de recollida d'informació: observació, entrevista, anàlisi de documents
- Triangulació d'observadors (observadors múltiples) o judici crític de companys<sup>16</sup>.
- Recollida de material d'adequació referencial: documents , gravacions, fotografies per a contrastar les informacions.
- Ús de descriptors de baixa inferència
- Establir una pista de revisió (és a dir, explicar bé com s'ha desenvolupat el procés)

---

<sup>15</sup> Aquestes estratègies s'empren per augmentar la credibilitat de la investigació naturalista, és a dir, controlar que la mesura obtinguda no estigui deformada pels instruments emprats, la subjectivitat dels informants o l'arbitrarietat de la interpretació (Bartolomé, 1995). Els criteris que permeten valorar-la són la credibilitat (validesa) , la transferibilitat (similar a la validesa externa, o grau de generalització), la dependència (consistència) i la confirmabilitat o l'objectivitat.

<sup>16</sup> Molts autors citen la triangulació com el pal de paller sobre el qual descansa el rigor científic de la investigació qualitativa. Triangular és un terme topogràfic que vol dir aconseguir aproximacions diverses d'una mateixa realitat per assegurar-nos que les nostres interpretacions, i les dades en les quals es basen, són correctes (Bartolomé, 1995:59).

- Auditoria de confirmabilitat (o d'objectivitat): porta un rèplica del procés per dos equips d'investigació separats (o avaluadors o observadors).

Com es pot observar, aquestes mesures permeten incrementar la seguretat que l'anàlisi, després de l'observació, entrevista, etc., es correspon a la realitat que s'intenta comprendre més que a les interpretacions o subjectivitat de l'avaluador. Es podrien resumir en dues estratègies: d'una banda, triangular (és a dir, aconseguir aproximacions diverses d'una mateixa realitat ja sigui al llarg del temps, d'observadors o d'instruments de recollida d'informació), i, de l'altra, deixar constància del procés per tal que es pugui resseguir, comprovar, mitjançant material referencial i descriptors de baixa inferència.

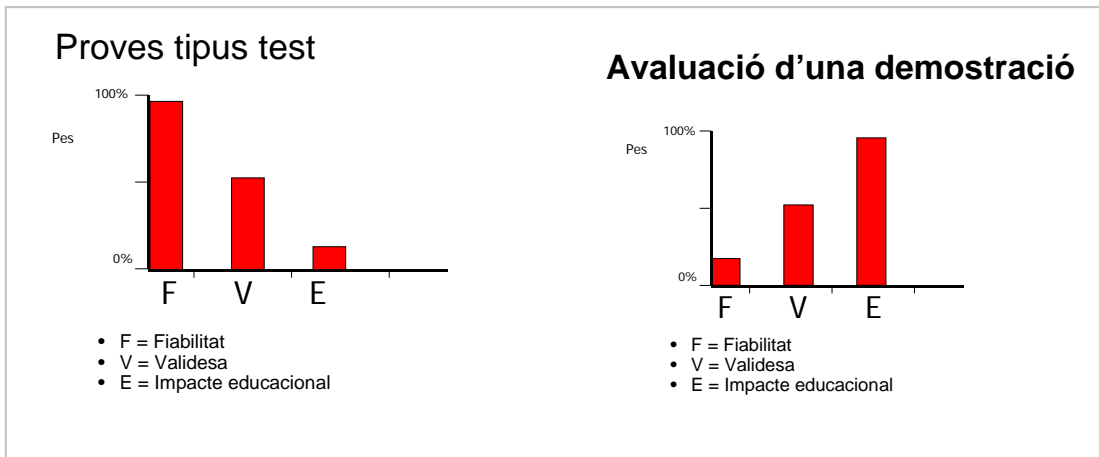
### **c) Altres criteris: l'impacte en l'aprenentatge i el cost**

Al primer apartat d'aquest capítol ja s'ha fet referència a la importància que té l'avaluació en l'enfocament de l'aprenentatge. Aspectes com el contingut a avaluar, el format que s'empra, però també la programació (calendari d'avaluació) o la normativa d'avaluació influeixen en l'enfocament que, finalment, l'estudiant adoptarà davant dels objectius d'aprenentatge, i, en conseqüència, en el nivell de competència finalment assolit. Les estratègies d'avaluació coherents amb el procés d'ensenyament aprenentatge tenen un impacte en l'aprenentatge major que les que no ho són. El criteri d'impacte en l'aprenentatge fa referència al grau en què l'avaluació fomenta un aproximació a l'aprenentatge adequada. Si desenvolupar competències implica, segons s'ha vist al Capítol 3, desenvolupar enfocaments profunds d'aprenentatge, caldrà evitar els que fomentin enfocaments superficials (la memorització), i emprar aquells que fomentin l'aprenentatge profund<sup>17</sup>.

---

<sup>17</sup> En aquest sentit, segons diversos estudis, els tests objectius recompensen estratègies d'aprenentatge superficial, mentre que els assajos, els dossiers d'aprenentatge, els projectes de recerca, l'autoavaluació, l'avaluació de companys, el *feedback* sobre l'aprenentatge fomenten enfocaments més profunds (Miller, Imrie i Cox, 1998).





Esquema 4.4.

L'Esquema 4.4 mostra dos dels diferents escenaris de combinació dels criteris de fiabilitat, validesa i impacte en l'aprenentatge. Així, per exemple, ens podríem imaginar una prova per certificar objectius de coneixement, tipus test, altament equitativa, amb una fiabilitat molt elevada, però una validesa més reduïda, i un impacte educacional, de motivació vers estratègies d'aprenentatge profund encara més baix; i un segon escenari, per exemple, una avaluació de pràctiques en una empresa, amb una fiabilitat molt discutible (avaluació pels observadors), una validesa relativa (coherent, pel que fa al context, però en perill a causa de la baixa fiabilitat), i un impacte educacional molt elevat (pel que representa de motivació, integració d'aprenentatges, transferència de coneixements acadèmics i la seva revalorització, etc.). No sembla desitjable la desaparició de cap dels dos escenaris, sinó que el més idoni és la seva combinació per tal de cobrir els avantatges diversos associats amb cadascuna de les estratègies avaluatives.

Finalment, un darrer criteri que sovint es menysté en els manuals d'avaluació, però que sens dubte afecta la presa de decisions, és el **cost** que té construir el mètode d'avaluació i de correcció, en relació als objectius que s'assoleixen, és a dir, la "practicitat" pel que fa a temes de temps i recursos de l'estratègia escollida.

## 4.4. Estratègies d'avaluació

---

Per tal com no hi ha correspondència entre cada prova i una determinada competència, s'analitzarà els potencials de diferents proves que d'entrada s'han classificat en dos grans grups: l'avaluació tradicionals, i l'avaluació d'execucions.

- L'avaluació tradicional: engloba el que podríem anomenar les típiques “proves de paper i llapis”, en la qual hi ha més èmfasi en els objectius de coneixements i saber. Dins l'avaluació tradicional hi ha proves que emfatitzen habilitats de baix ordre (record, comprensió), mentre que n'hi ha d'altres que emfatitzen el pensament d'alt ordre (aplicació, síntesi, avaluació).
- L'avaluació d'execucions és, com es veurà, molt variada, i permet abastar un rang molt més ampli de competències, ja sigui d'habilitats disciplinars (que els graduats d'infermeria sàpiguen posar una injecció, o els de farmàcia pipetejar), com de competències transversals (comunicació oral, pensament crític, etc.).

Mentre que els criteris de coherència i teòrica validesa són els principals arguments que s'esgrimeixen per introduir noves metodologies d'avaluació, la baixa fiabilitat (i, de retruc, validesa) de les proves d'execució és el criteri més emprat de resistència al canvi.

Si bé els dos tipus d'estratègies permeten tant l'avaluació referida a la norma com la criterial, en l'avaluació tradicional és més fiable fer comparacions referides a la norma que en la d'execucions, i, per contra, en la d'execucions és més factible fer una avaluació criterial dels perfils de competències, perquè és més variada.

### a) L'avaluació tradicional

Emfatitza més els objectius de coneixements, de saber. Però com s'ha vist al Capítol 3, hi ha diversos nivells de profunditat en el domini d'aquests coneixements: del record, a la comprensió, fins a l'ús del record per a l'anàlisi, síntesi i avaluació. Dins l'avaluació tradicional, analitzarem els tests objectius, les proves lliures escrites, i les preguntes de problemes científicomatemàtiques.

## Quadre 4.1<sup>18</sup>. Proves tradicionals: característiques, utilitat, fiabilitat i validesa, i consells per un bon disseny

	Característiques	Útils per mesurar	Fiabilitat i Validesa	Per un bon disseny...
Tests objectius	<p>Són proves on es requereix seleccionar la resposta correcta d'un conjunt de respostes possibles ítems de vertader/fals, ítems d'emparal·lel·lisme, d'elecció múltiple..).</p> <p>Els ítems poden ser gràfics, text, exemples, o fins i tot casos.</p> <p>Un cop construïts són de fàcil aplicació i correcció, permeten un retorn o <i>feedback</i> a l'estudiant ràpid</p>	<p>Objectius com ara els de reconèixer - discriminar informació, aplicació de principis o regles i interpretació de dades</p> <p>Reforça més el pensament selectiu que els processos mentals dirigits a la construcció del coneixement</p>	<p>Fiabilitat: La puntuació de la prova es veu menys alterada per factors estranys al procés de puntuació.</p> <p>Permeten aplicar un judici valoratiu amb el mateix criteri a diferents execucions, mentre que en les subjectives no es pot assegurar la igualtat del criteri. La qualificació <i>d'objectius</i> fa referència a les condicions d'aplicació de la prova i al tractament i interpretació dels resultats, però no indica que siguin més objectives pel que fa al punt de vista d'una millor valoració del rendiment dels estudiants</p> <p>Validesa: Permeten avaluar un ampli ventall de contingut, la qual cosa n'augmenta la validesa. La validesa es pot millorar a través de l'anàlisi del funcionament dels ítems</p>	<p>Cal dissenyar adequadament els ítems per tal que no permetin només mesurar la capacitat de record.</p>
Preguntes curtes	<p>Conjunt de preguntes obertes on l'alumnat elabora i estructura la seva resposta amb total llibertat.</p> <p>Els ítems poden ser gràfics, text, exemples, o fins i tot casos que requereixen l'elicitació o elaboració d'una resposta.</p> <p>Segons l'amplitud de resposta que s'exigeix es diferencia entre proves d'assaig ampli o desenvolupament de temes, i proves d'assaig restringit o de resposta curta.</p> <p>És més ràpid de construir que els test objectius i és fàcil i barat d'administrar.</p>	<p>Tant poden implicar habilitats cognitives d'alt ordre (transferència i integració de l'aprenentatge), com la simple repetició d'un contingut prèviament memoritzat. Té, però, potencial per mostrar l'aprenentatge profund, ja que es requereix la construcció de la resposta.</p> <p>Són pertinents per avaluar objectius referits a: evocació de la informació, interpretació de l'evidència, construcció d'un disseny, generació d'hipòtesis, exposar la informació per a una decisió o explicitar les fases d'un procés</p> <p>Permeten valorar l'ús del vocabulari i raonament conceptual propi d'una àrea de coneixements</p>	<p>Les seves propietats psicomètriques són qüestionables (dificultats en la fiabilitat inter-avaluadors, cobriment de dominis restringits de coneixement)</p> <p>Les proves de preguntes curtes són més útils per avaluar un repertori adequat dels continguts de l'aprenentatge que les proves d'assaig ampli</p> <p>Pot desavantatjar els estudiants amb habilitats de comunicació escrita més pobres</p>	<p>Cal:</p> <ul style="list-style-type: none"> <li>* Vigilar que les preguntes no requereixen respostes purament memorístiques</li> <li>* Clarificar les expectatives d'execució especialment a les instruccions</li> <li>* Disposar d'una taula o esquema de correcció el màxim de detallada possible, que permeti millorar-ne la fiabilitat i donar un <i>feedback</i> fiable als estudiants</li> </ul>

<sup>18</sup> Font: síntesi personal a partir de les fonts següents: APA (2003), Bordas, Cabrera (2001), Mateo (2000), Miller, Imrie i Cox (2001), Palomba i Banta (1999); Rodríguez Espinar (2000). Per les qüestions relatives als diferents tipus de problemes i dificultats en la transferència de la competència de resolució de problemes, s'ha emprat Garnham i Oakhill (1996).

	Característiques	Útils per mesurar	Fiabilitat i Validesa	Per un bon disseny...
Proves científicomatemàtiques	<p>Estan a mig camí entre les proves de format lliure i les de format objectiu, per tal com exigeixen la construcció de la resposta, però permeten una correcció més objectiva<sup>19</sup></p> <p>La complexitat de problemes pot variar en funció del nombre de passos per a la seva resolució, el grau d'abstracció que impliquen, i per les operacions cognitives implicades</p> <p>El grau de la novetat influirà en la dificultat del problema: així, és més fàcil recórrer a una analogia si hi ha similituds tan superficials com estructurals entre els problemes</p>	<p>Són una bona forma de comprovar la comprensió i aplicació (en principi), en contrast amb la memorització. Rellevants per la dimensió tecnicoprofessional.</p> <p>Els problemes, com els assaigs permeten veure el desenvolupament de certes competències transversals, com el pensament crític, la presa de decisions</p> <p>Cal diferenciar entre l'<b>aplicació simple</b> i la <b>resolució de problemes</b>: reconeixement o record de la informació <i>versus</i> la seva reestructuració o reelaboració, i grau en què els exercicis són rutinaris, <i>versus</i> grau en què són originals.</p>	<p>Bona fiabilitat (tot i que també cal tenir clars els criteris de correcció) i validesa (poden abastar un rang ampli de continguts).</p> <p>Pel que fa a la validesa, cal tenir en compte qüestions sobre la transferència de la competència de resolució de problemes, sembla apuntar que l'habilitat és transferible però dins el mateix domini (Garnham i Oakhill, 1996)<sup>20</sup>.</p>	<p>En el disseny dels problemes, cal ser conscient de la dificultat de transferència entre dominis en el continu:</p> <ul style="list-style-type: none"> <li>- Problemes que requereixen aplicar una rutina (memorització)</li> <li>- Problemes que requereixen una reestructuració, i l'enunciat presenta similitud superficial i profunda amb altres problemes resolts</li> <li>- Reestructuració a partir de similitud profunda, però no superficial</li> </ul>

<sup>19</sup> Si bé el professor Antoni Sans (2004) mostra com la correcció d'un examen de matemàtiques pot portar a puntuacions molt diverses en funció dels criteris de qui corregeixi: per exemple, si es tenen en compte els errors en una operació que no era requerida, l'atribució d'un error (distracció vs error conceptual), etc.

<sup>20</sup> La recerca d'Staver & Lumpe (1995, citat a Bodner i Herron, 2002) assenyalava com una de les causes de fracàs en la transferència de conceptes entre diferents assignatures d'un mateix ensenyament l'ús d'algoritmes o regles memoritzades, juntament amb la insuficient comprensió dels conceptes implicats i la baixa habilitat per transferir la comprensió entre nivells atòmic/moleculars i els macroscòpics en la resolució de problemes (p 249).

## b) L'avaluació d'execucions

L'avaluació d'execucions és el procés d'emprar les activitats dels estudiants o productes que han elaborat, per avaluar el coneixement, habilitats i desenvolupament. Aquest tipus d'avaluació es caracteritza per intentar simular com sigui més possible una situació real que la persona ha d'afrontar. Consisteix a requerir als estudiants que demostrin, construeixin o desenvolupin un producte o solució a partir d'unes determinades condicions i estàndards, és a dir, proves en les quals cal una construcció de la resposta (Rodríguez, 2000). Aquests mètodes inclouen assajos, presentacions orals, exhibicions i demostracions (Palomba; Banta, 1999). No es tracta d'una invenció recent, sempre s'han emprat aquest tipus de proves, la diferència rau en la importància i el grau d'extensivitat que ara s'està donant.

Normalment són situacions que posen en joc més d'una capacitat de la persona. En ocasions acaba un producte (un dibuix, una peça, resoldre un cas, etc.), en altres, l'important és valorar l'actuació durant la seva realització (Bordas, Cabrera, 2001). Típicament aquestes avaluacions requereixen demostrar el pensament crític, raonament creatiu i solució de problemes, emprant informació en un context significatiu.

Un dels grans desafiaments de l'avaluació d'execucions és l'elaboració d'aproximacions profitoses per a avaluar les activitats o els productes generats pels estudiants (Palomba i Banta, 1999). Bordas i Cabrera (2001) assenyalen una sèrie de criteris per a elaborar aquest tipus de proves:

- Les normes a les quals ha d'ajustar-se el producte han d'establir-se amb claredat
- S'ha de proporcionar a les persones que hagin de fer la prova els criteris d'avaluació que s'utilitzaran
- En la valoració de processos cal especificar els comportaments que seran objecte de valoració
- En la valoració de productes: desenvolupar un esquema, índex o llistats de continguts que han d'estar presents en el producte i els criteris que es tindran en compte per a la seva valoració
- Establir el pes relatiu dels diferents aspectes en una valoració global

- Des de l'inici s'han d'especificar amb claredat les *regles del joc* a les quals la persona ha d'ajustar la seva actuació o el seu producte
- Procurar l'objectivitat en el registre de les actuacions o en l'anàlisi del producte mitjançant algun instrument d'observació o d'anàlisi

Quines competències promouen? El grau en què es desenvolupin altres competències transversals dependrà del tipus de prova. A la taula següent analitzarem les següents:

- Proves orals
- Productes / Execucions: productes escrits, projectes de recerca, treball en grup, estudis de cas
- Pràctiques estructurades
- Avaluació de laboratori
- Autoavaluació
- Dossier d'aprenentatge
- El treball en equip

## Quadre 4.2<sup>21</sup>. L'avaluació d'execucions: característiques, utilitat, fiabilitat i validesa, i consells per un bon disseny

	Característiques	Útils per mesurar	Fiabilitat i Validesa	Per un bon disseny...
Proves orals	<p>Tradicionalment impliquen un o dos examinadors que fan qüestions als estudiants sobre la comprensió i habilitat d'aplicar el que han après, però també s'inclouen debats, <i>role-playing</i>, etc.</p>	<p>Permeten valorar la <b>capacitat de comunicació i habilitats interactives</b>, habilitats que no es poden avaluar d'altra manera, i, a més a més, promouen el pensament autònom mitjançant l'estructura qüestió - resposta. L'avaluació és, a més, una oportunitat per a posar en pràctica l'expressió oral, millorant, per tant, aquestes habilitats</p>	<p>El principal inconvenient és que permeten considerable llibertat a l'examinador per variar les qüestions als estudiants i són difícils de qualificar, cosa que les converteix en poc fiables.</p> <p>Són les proves més adequades (coherents) per valorar la competència de comunicació oral. Tanmateix, la capacitat oral en si no acostuma a ser objecte d'avaluació en les proves orals, sinó que tan sols s'avalua el coneixement acadèmic.</p> <p>De fet, algunes estudis han mostrat que la majoria de preguntes requereixen poc més que el record d'alguns fragments d'informació, cosa que es pot avaluar més fàcilment i de manera més fiable amb tests escrits objectius.</p> <p>Desafavoreixen els estudiants amb por de parlar en públic.</p>	<p>Les mesures per incrementar la fiabilitat passen per assegurar que els professors preguntin el mateix (equivalent en contingut i dificultat), preparar fulls de puntuació, fer servir tants exàmens com sigui possible, que cada estudiant tingui el mateix temps per respondre, establir controls per possibles biaixos per edat, sexe o ètnia, etc</p> <p>A més a més, caldrà vèncer la resistència dels estudiants a parlar en públic, i assegurar la qualitat de les presentacions mitjançant guies, desenvolupament de criteris d'execució, etc</p>

<sup>21</sup> Font: síntesi personal a partir de les fonts següents: APA (2003), Bordas, Cabrera (2001), Mateo (2000), Miller, Imrie i Cox (2001), Palomba i Banta (1999); Rodríguez Espinar (2000). Per les qüestions relatives als diferents tipus de problemes i dificultats en la transferència de la competència de resolució de problemes s'ha emprat Garnham i Oakhill (1996).

	Característiques	Útils per mesurar	Fiabilitat i Validesa	Per un bon disseny...
Productes	<p>Són <b>específics per ensenyaments</b>, per exemple articles de diari per estudiants de periodisme, quadres per estudiants de belles arts, mapes pels de geografia, programes informàtics pels d'informàtica, etc.</p> <p>A banda de productes, però, l'avaluació d'execucions o del rendiment es pot emprar per avaluar <b>demostracions</b> del treball de l'estudiant: fer servir un instrument, dur a terme una entrevista, etc. Es pot observar infermers, futurs mestres conduint una classe, o els estudiants al laboratori. També es poden emprar programes de <b>simulació</b>. Aquestes execucions solen donar més informació directa sobre l'aprenentatge que els tests objectius</p> <p>El principal inconvenient d'aquesta avaluació és el temps de correcció</p> <p>És difícil de construir i mesurar</p>	<p>Eina ideal per avaluar <b>competències disciplinars o tècniques</b> pròpies de l'àrea de coneixement. Promou la transferència dels coneixements acadèmics i que s'afavoreixen habilitats cognitives d'alt ordre</p> <p>Cal afegir, com avantatge per al procés d'aprenentatge, la motivació que suposa pels estudiants una situació d'avaluació realista</p> <p>El grau en què es desenvolupin altres competències transversals dependrà del tipus de prova (productes escrits, gràfics, pòsters, estudis de cas...). Per exemple:</p> <p><b>Projectes de recerca:</b> manera d'avaluar la capacitat de gestió de la informació, l'aplicació dels coneixements i competències disciplinars en la resolució de problemes. Situats al final del currículum, motiva els estudiants des del principi del seu recorregut acadèmic, fomentant la responsabilitat de l'estudiant, i la coherència del programa</p> <p><b>Pòsters:</b> Donen l'oportunitat per integrar les competències de comunicació (oral, escrita, gràfica) amb continguts acadèmics</p> <p><b>Estudis de cas i longitudinals:</b> són una altra modalitat de resolució de problemes, en la qual destaca la riquesa de detalls.</p>	<p>És difícil de construir (l'elecció de la mostra condiciona la validesa) i mesurar (subjectivitat – fiabilitat de la correcció)</p> <p>Hi ha el perill que els estudiants es basin més en el sentit comú que en els seus coneixements en situació de pressió.</p> <p>Un altre element que afectaria la seva validesa és el perill de plagi</p> <p>Segons les proves, per exemple els estudis de cas, o els grans problemes, en ser una mostra petita de contingut (això sí, amb profunditat), es corre el risc de limitar la generalització i ometre, per tant, la transferència del coneixement</p>	<p>Per tal d'afavorir la transferència cal que la situació d'avaluació es correspongui estretament amb les condicions d'aprenentatge.</p> <p>La qualitat de les instruccions, els criteris avaluatius clarament definits i la formació dels avaluadors influiran en la seva validesa</p> <p>Pot resultar útils emprar instruments com els <i>cheklists</i><sup>22</sup>, fulls de valoració<sup>23</sup>, o emprar tècniques de la metodologia qualitativa de recerca per incrementar el valor científic de l'avaluació (triangulació, descriptors de baixa inferència, seguiment de ruta...</p>

<sup>22</sup> Un *checklist* és una escala de dos punts (la fiabilitat baixa si l'escala és més gran de quatre punts). Cal decidir si cada component de la llista és present/absent, adequat/inadequat, satisfactori/insatisfactori. És important que tots els components sigui clarament observables.

<sup>23</sup> Per assegurar la fiabilitat d'aquests full de valoració, cal emprar múltiples valoracions independents del mateix estudiant sobre la mateixa activitat.



	Característiques	Útils per mesurar	Fiabilitat i Validesa	Per un bon disseny...
Pràctiques estructurades	<p>Són uns tipus de proves d'execucions. Consisteixen en exàmens pràctics estructurats objectivament, tenen per objectiu provar un ampli ventall d'habilitats d'una forma objectiva</p> <p>Els estudiants passen per una sèrie d'estacions i duen a terme una varietat de tasques practiques</p> <p>Aquesta aproximació, inicialment desenvolupada com a part integral dels exàmens mèdics, ha estat desenvolupada i adoptada posteriorment per una varietat de professions.</p>	Competències disciplinar específiques o tècniques	<p>Bona fiabilitat, a costa d'un cost elevat (multiobservadors).</p> <p>Bona validesa per l'autenticitat de les situacions d'avaluació (s'assegura la transferència)</p>	Fulls de puntuació i <i>checklists</i> es preparen abans per millorar la fiabilitat de a puntuació. Els estudiants s'examinen del mateix contingut i es puntuen amb el mateix criteri pels mateixos examinadors
Av. laboratori	<p>És un tipus de proves d'execucions. L'avaluació de laboratori té lloc en un entorn realista i requereix la complementació d'una tasca real. L'avaluació de l'execució pot ser sobre el procés, el producte o ambdós elements.</p>	Competències de laboratori <sup>24</sup>	Massa sovint, l'avaluació es basa per complet en un informe escrit i no en l'observació directa de l'execució dels estudiants, això produeix un desajust entre els objectius establerts i el focus d'avaluació. L'observació presenta dificultats en la qualificació a causa de la subjectivitat de l'avaluador.	Definir clarament els objectius i resultats esperats de l'execució, així com un calendari per l'avaluació, minimitzarà les inconsistències entre observadors <sup>25</sup> .
L'auto-avaluació	L'autoavaluació es pot definir com l'habilitat de ser capaços de fer judicis acurats sobre el seu rendiment i les seves fortaleces i debilitats.	Es fa servir per millorar la comprensió sobre la pròpia habilitat i l'execució val la pena educacionalment, i anima l'obertura i a la honestat sobre l'avaluació. Aquesta capacitat metacognitiva, és fonamental per aprendre noves competències	Els principals inconvenients que presenten és que és propensa a biaixos dels estudiants, ja sigui per sobre o subautoestima. Entre els factors d'error de les autoavaluacions trobem la falsificació intencionada, l'acceptació social o l'aquiescència <sup>26</sup> .	L'establiment dels criteris i estàndards, pas previ ineludible de les autoavaluacions. Una alternativa és contrastar la pròpia marca amb la que dona un company. El professor també pot puntuar una mostra aleatòria per establir controls i desanimar l'engany o la desil·lusió <sup>27</sup> .

<sup>24</sup> Per Bennet i O'Neale (1998), formarien part d'aquestes competències l'observació, la manipulació, la interpretació, les competències tècniques (cromatografia, espectrografia, precipitació) i el disseny expert.

<sup>25</sup> Per a veure graelles possibles de valoració consultar: Miller, Imrie i Cox (1998), Bennet i O'Neale (1998)

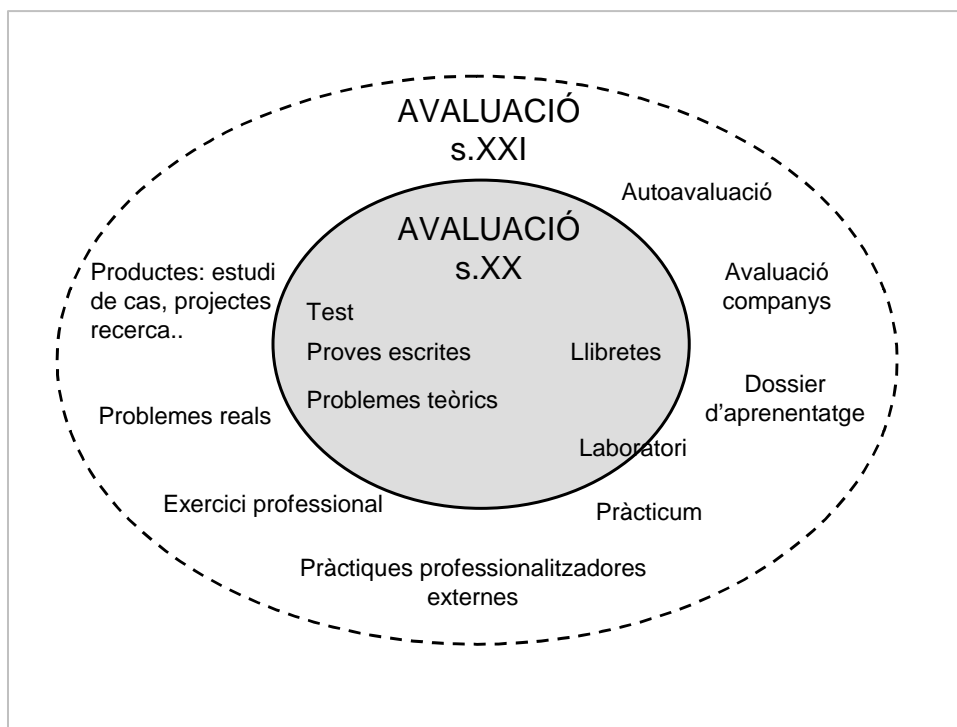
<sup>26</sup> Això fa que l'autoavaluació no sigui un mètode adequat per a la investigació de certs aspectes, per exemple, de la personalitat: McLelland va citar-la com la calamitat més gran de la investigació de la personalitat (citada a Amelang i Bartusek, 1991). Per què, doncs, s'ha d'emprar en entorns formatius? Si estem en un context on és necessari l'aprenentatge al llarg de la vida, aleshores la capacitat de ser conscients de les debilitats i fortaleces del nostre rendiment és un element indispensable.

	Característiques	Útils per mesurar	Fiabilitat i Validesa	Per un bon disseny...
d'aprenentatge	<p>Els dossiers d'aprenentatge són una col·lecció selectiva, deliberada i validada dels treballs fets per l'estudiant en què es reflecteixen els esforços, els progressos i aprenentatges en una àrea específica al llarg d'un període de temps.</p> <p>Els estudiants ajunten, presenten, expliquen i <b>avaluen</b> el seu aprenentatge en relació als objectius del curs i als seus propis objectius o expectatives.</p> <p>Consumeix temps i és difícil d'avaluar, el contingut variarà àmpliament entre els estudiants</p>	<p>La seva finalitat, <b>és un balanç del progrés i del desenvolupament</b> dels aprenentatges de l'estudiant</p> <p>Afavoreix el desenvolupament de competències d'independència o autonomia, reflexió i autoorientació</p> <p>Promou l'autoconsciència, la responsabilitat sobre el seu aprenentatge,</p> <p>Il·lustra tendències longitudinals, subratlla les fortaleses de l'aprenentatge i identifica les debilitats a millorar.</p>	<p>És coherent amb l'enfocament de l'aprenentatge centrat en l'estudiant</p> <p>La validesa dels dossiers en relació a la competència de reflexió, o metacognitiva és clara en aquesta situació, però la seva fiabilitat per a avaluacions sumatives encara s'ha de determinar</p>	<p>Un dossier d'aprenentatge ha de tenir les parts següents: una introducció on es defineixen els objectius i què espera l'estudiant satisfer; la presentació de les peces d'evidència que demostrin que hi ha hagut un aprenentatge; i, finalment, una explicació de per què s'han triat i presentat aquests elements i una avaluació dels resultats d'aprenentatge</p>
Treball en equip	<p>Aquesta estratègia reparteix la responsabilitat de l'aprenentatge en grups. S'espera dels estudiants que interactuïn, comparteixin idees i recursos, i que es fomenti l'aprenentatge entre col·legues assumint mútua responsabilitat per l'assoliment de l'aprenentatge.</p> <p>És especialment útil quan les fonts o recursos d'aprenentatge estan disponibles, i hi ha una gran quantitat de material a assimilar</p>	<p>A més de posar en pràctica les habilitats interactives (comunicació, treball en equip, lideratge) incrementa la motivació, facilita l'aprenentatge col·laboratiu i permet que l'estudiant s'enfronti a problemes molt més complexos que els de resolució individual</p> <p>També es posa en joc la gestió de temps, i es promou el treball independent a un nivell més profund</p> <p>Promou la implicació, ja que acostumen a simular activitats reals</p>	<p>Els <i>ganduls socials</i> poden imposar judicis equitatius sobre la puntuació, el fet que idees errònies que no són copsades i corregides es dispersen entre membres dels grups. A més a més, consumeix temps, i pot resultar difícil pel professorat jutjar quan cal redirigir o rescatar els grups d'estudiants a temps</p> <p>És un desafiament avaluar els resultats de l'aprenentatge en grup de manera que sigui <i>fair</i> pels individus, però que reconegui les dinàmiques i realitat particulars de l'aprenentatge en grup</p>	<p>Cal dissenyar bé la demanda de manera que sigui motivadora i que per resoldre-la calgui l'esforç de tots els membres</p> <p>El pes avaluatiu ha de ser proporcional a l'esforç implicat.</p> <p>Cal penalitzar fortament el plagi, i evitar-lo no demanant la reproducció o el record de conjunts de fets, teories, conceptes, etc.</p> <p>Altres formes d'optimitzar la valoració inclourien: ús d'una nota global matisada per l'avaluació de companys (a partir d'uns criteris escrits) i entrevistes complementàries amb els estudiants (veure Cannon i Newble, 2000)</p>

<sup>27</sup> La investigació de Boud (citada a Cannon i Newble, 2000) demostra que si l'esquema d'avaluació està ben dissenyat i els estudiants es puntuen en relació a l'assoliment (i no en l'esforç), les notes seran consistents amb les del professorat

### c) Conclusions

L'avaluació és una fase en el procés d'ensenyament aprenentatge i cal dissenyar-la en conseqüència, *com si l'aprenentatge fos allò més essencial*. Per fer-ho hem de combinar els criteris de coherència, validesa i consistència, i impacte educacional, a la vegada que es prendran en consideració qüestions relatives al cost. En ampliar els objectius de formació de coneixements a competències, de la cognició a l'acció reflexiva, cal diversificar les estratègies avaluatives. Hi ha la tendència d'introduir proves d'avaluació d'execucions, perquè en ser més variades permeten abastar competències més àmplies i, per tant, fer avaluacions criterials per aquest rang; a més, afavoreixen enfocaments "actius" d'aprenentatge, d'aplicació del coneixement. L'esquema 4.5 vol il·lustrar que les noves estratègies avaluatives s'afegeixen a les tradicionals; enriqueixen les mostres d'aprenentatge i afavoreixen aquest escenari on s'aprofiten els avantatges d'una pluralitat de fonts d'avaluació.



Esquema 4.5.

Les diferents proves de l'avaluació tradicional i les proves d'avaluació d'execucions difereixen en la seva potencialitat. El Quadre 4.3 ofereix una comparació sintètica d'aquestes estratègies a partir de diversos criteris:

**Quadre 4.3. Resum dels avantatges i desavantatges de les formes d'avaluació**

Criteris	Proves tradicionals				Proves d'execució
	Proves objectives	Proves preguntes curtes	Proves preguntes curtes (problemes)	Proves d'assaig ampli	Productes, demostracions, treballs en equip, dossiers d'aprenentatge...
1. Visió comprensiva dels coneixements (fets, conceptes, etc.)	++	+	+	-	++ (integració) - (repertori)
2. Per distingir diferències precises entre les persones	++	-	-	-	-
3. Per valorar l'ús del vocabulari i raonament conceptual propi d'una àrea de coneixement	-	+	- (vocabulari) + (raonament)	++	++ (informe/exposició/treball...menys si es tracta d'una demostració: avaluació de laboratori...)
4. Per valorar l'habilitat d'organitzar, integrar, sintetitzar i argumentar	-	+	+ (organitzar) - (sintetitzar)	++	++ (en treballs, projectes, elaboració d'articles per revistes d'estudiants...)
5. Per valorar enfocament originals als problemes	-	+	+	++	+
6. Per valorar habilitats cognitives d'alt ordre	-	+	+	+	+
7. Cost en temps	++	+	+	-	--
8. Valor pels estudiants	-	+	+	+	++
9. Informació pel programa	+	+	+	-	++
10. Per valorar competències tècniques/específiques	-	-	+	+	++ (avaluació de laboratori, demostracions, pràctiques a empreses, projectes..)
11. Per valorar habilitat metacognitives	-	-	-	-	+ (els dossiers d'aprenentatge, autoavaluacions, avaluació de companys)
12. Objectivitat en la qualificació	++	-	+	-	-

**Nota:** Les proves objectives, per estar ben construïdes, tenen un cost de temps molt elevat d'elaboració, per poc de correcció (passa al revés amb les proves curtes). Pel que fa a valorar les habilitats d'alt ordre dependrà de si les proves objectives estan ben construïdes. Com que això, segons la literatura consultada, no és gaire comú, s'afirma que no valoren les habilitats cognitives d'alt ordre, que afavoreixen la memorització i que són ràpides de corregir. La taula reflecteix, doncs, aquesta la situació, però no la potencialitat de les proves objectives

Una altra manera de comparar les proves és per la seva coherència en relació als objectius d'ensenyament- aprenentatge, el Quadre 4.4 presenta, per diferents tipus de competències

transversals, una proposta de proves avaluatives coherents o susceptibles d'oferir informació sobre el grau en què s'ha assolit un determinat desenvolupament:

#### Quadre 4.4. Competències transversals i proves susceptibles de mesurar-les

<b>Coneixements disciplinars</b>	Proves objectives, preguntes obertes curtes, Temes, Problemes científicomatemàtics
<b>Competències disciplinars específiques</b>	Avaluació de productes (pòsters). Avaluació d'execucions ( <i>checklist</i> , graelles valoratives, laboratori, pràctiques a empreses, <i>assessment center</i> <sup>28</sup> ). Projectes de recerca
<b>Pensament crític</b>	Projectes de recerca, estudis de cas, Simulacions, Dossiers d'aprenentatge, Avaluació al laboratori
<b>Resolució de problemes</b>	Col·lecció de problemes, Treball en grup, Tesina/Projecte, avaluació al laboratori..
<b>Expressió escrita</b>	Treballs de curs, dossiers (assajos) d'aprenentatge, informes de recerca, llibreta laboratori
<b>Comunicació oral</b>	Presentacions, Debats, Entrevistes simulades, Respostes – Preguntes laboratori
<b>Treball en equip</b>	Treball en grup, Avaluació entre iguals, Autoavaluació
<b>Documentació</b>	Treballs bibliogràfics, Projectes de recerca, Estudi de casos...
<b>Raonament ètic i valors</b>	Indirectament a través d'execucions (Projecte final de carrera, pràcticum, estudis de casos, etc.). O bé més directament a través d'instruments de mesura qualitatiu (debats, entrevistes, focus grups amb estudiants, etc.), o quantitativus (escales d'actituds, enquestes)
<b>Aprenentatge al llarg de la vida</b>	Autoavaluacions, avaluació dels companys, dossier d'aprenentatge
<b>Iniciativa</b>	<i>In-basket</i> <sup>29</sup> , avaluació d'execucions, Autoavaluació, pràcticum...

<sup>28</sup> Els *assessment center* són similars a les Pràctiques estructurades descrites a la Taula 2. S'identifiquen aspectes claus de l'aprenentatge i s'empren simulacions i altres tècniques per generar comportaments als estudiants. Dos o més avaluadors observen i avaluen els comportaments

<sup>29</sup> En els *in-basket*, els estudiants fan el rol de un supervisor: manegen cartes, *memos* i informe continguts en l'*in-basket*

## 4.5. La gestió de l'avaluació

---

Tota la complexitat que s'ha reflectit abans posa de manifest que l'avaluació ja no pot ser una qüestió entre el professor/l'assignatura i els seus estudiants, i passa a ser una qüestió institucional.

Igual que amb el concepte de competències, la introducció de nous conceptes en l'avaluació genera canvis o implicacions a nivell cultural (concepcions sobre els fins de l'avaluació), tecnològiques (nous instruments per avaluar), i politico-socials (nous rols en l'avaluació per part d'estudiants i professorat, canvi d'estructures d'avaluació, com, per exemple, una major presència de juntes d'avaluació, etc.).

Fins i tot amb professorat amb una visió de l'avaluació innovadora (com quelcom que ha d'anar molt més enllà de l'avaluació sumativa, la importància de l'avaluació en l'aprenentatge, etc.), el pes de la inèrcia o el pes de la tradició pot fer que res canviï, i es mantingui l'èmfasi amb formes d'avaluació tradicionals i formals, com es va constatar en la investigació de Bennet, Dune i Carré (2000)<sup>30</sup>.

### El canvi de les pràctiques avaluatives

És possible canviar les pràctiques avaluatives existents? Què és més fàcil, es poden canviar les pràctiques avaluatives sense canviar les concepcions de l'avaluació? Les concepcions que té la comunitat acadèmica del coneixement de la disciplina determinaran la forma en la qual es creu que es fomenta l'aprenentatge i el tipus de demanda avaluativa (Samuelowicz i Bain, 2001). Així, si es concep el coneixement com quelcom extern als estudiants, com un col·lecció de bits importants (conceptes, teories) elaborats per experts en el camp, i que l'aprenentatge consisteix a aplicar allò ja establert en problemes determinats, les pràctiques docents fomentaran l'adquisició d'aquests *bits* d'informació (*teacher oriented*) i les avaluatives la seva *reproducció* en el context adequat. En canvi, la visió del coneixement com quelcom que ha de ser internalitzat, reorganitzat per l'aprenent comporta metodologies en les qual es requereixi

---

<sup>30</sup> Els autors d'aquesta recerca ho atribueixen al fet que perquè canviïn les formes d'avaluació, cal canviar també la concepció dels objectius de l'educació que té el professorat. Així, una aproximació tradicional epistemològica assumeix que l'aprenentatge és un increment quantitatiu de coneixement, i que la comprensió s'incrementarà amb la proporció de la quantitat de fets apresos (Ramsden 1992, citat a Bennet, Dune i Carré, 2000). Per tant, ensenyar és fer servir mètodes per incrementar el coneixement i involucrar els estudiants en mètodes de recerca o el procés de generació de la ciència seria una pèrdua de temps. En aquest context, l'avaluació és la mesura de l'increment de coneixement.

que l'estudiant integri informació de fonts diverses (*student centred*) i tasques d'avaluació en les quals els estudiants apliquin la seva pròpia comprensió a problemes oberts. Això implica que els mètodes d'avaluació "transformadors" obtindran molta resistència, no per arguments molt més profunds que el nombre d'alumnes per classe o el poc suport docent, sinó a causa de la concepció del coneixement en la disciplina (Samuelowicz i Bain, 2002)<sup>31</sup>. No es poden, per tant, canviar d'un dia per l'altre totes les estratègies avaluatives, ara bé, la introducció gradual d'estratègies cada cop més diverses permetria anar vençant resistències i iniciar el canvi de les concepcions sobre l'aprenentatge.

Diversificació dels mètodes avaluatius, avaluacions formatives, autoavaluacions... tot això implica un canvi en tot el sistema. Com gestionar aquest canvi?

- Adoptant un enfocament sistemàtic de l'avaluació dels aprenentatges de manera que s'assoleixi que el binomi cost benefici sigui clarament favorable pel segon terme, considerant professorat, alumnat i institució (Brown i Knight, 1994).
- Tenir en compte el context en el qual s'insereixen no només els objectius de l'assignatura o més enllà del perfil de formació (dimensió intrínseca). El model d'avaluació que s'adopta no és aliè al model organitzatiu del desenvolupament de l'ensenyament i de l'avaluació (dimensió extrínseca), la semestralitat o trimestralitat, el nombre d'alumnes o de professors, la normativa d'exàmens, etc. són elements que condicionen l'avaluació. Finalment, cal prendre en consideració la dimensió interactiva: l'estil docent i l'estil d'aprenentatge (Prat i Rodríguez, 2004).

---

<sup>31</sup> En les dues investigacions citades (Samuelowicz i Bain, 2001 i 2002) la Química apareix com a exemple on el coneixement és quelcom extern a l'estudiant, i el procés d'ensenyament aprenentatge com un mitjà per a assolir-lo. Així, un dels quatre professors de química entrevistats, afirma que existint un corpus de coneixement definit per la disciplina i desenvolupat per grans pensadors "seria absurd" esperar que els estudiants "reinventessin la roda" (Samuelowicz i Bain, 2001:313). Aplicar aquest plantejament en el marc del paradigma impulsat per la convergència europea (l'estudiant com a centre d'aprenentatge, metodologies actives d'aprenentatge com ara el descobriment..) comporten arribar a la conclusió que l'enfocament europeu potser és més adequat per unes disciplines que per unes altres. Ara bé, en l'estudi esmentat només s'entrevisten quatre professors de química, i no s'esmenta si d'un sol departament o només d'un, la qual cosa és una limitació important per a generalitzar els resultats en tota una disciplina. A més a més, probablement tot depèn de la rotunditat amb la qual s'interpreti que els estudiants han de descobrir el coneixement: des del principi (tal i com van fer els científics de l'època) o bé fent les connexions finals, hi ha, doncs, un continuum de descobriment o argumentació.

En el mateix sentit, un estudi de Sheppard i Gilbert (1991, citat a Cliff i Woodward, 2004) mostra que els estudiants de l'àmbit de ciències és més probable que creguin que el coneixement és inviolable i incontestable que els estudiants d'Humanitats. Les concepcions epistemològiques tenen, segons Cliff i Woodward, implicacions en la forma de concebre l'aprenentatge: així per un estudiant d'Humanitats l'aprenentatge més que consistir a memoritzar, consisteix a interpretar, mentre que si es concep l'aprenentatge com aprendre fets absolut i incontestables, l'estratègia més adient és la memorització.

- El **treball mancomunat del professorat** no es pot veure com una acció individual (Mateo, 2001). De la mateixa manera que les assignatures no donen els mateixos continguts, tampoc es pot pretendre que s'hi desenvolupin totes les competències en cada assignatura i totes les formes d'avaluació en cadascuna d'elles. Només es podrà afrontar el repte que ens demana l'Espai europeu d'educació superior si tenim en compte la totalitat d'assignatures que convergeixen durant un mateix període incidint en l'aprenentatge de l'estudiant (Rodríguez, 2000). El disseny de l'aprenentatge, i consegüentment el de l'avaluació, s'ha de fer des de la perspectiva d'una comunitat de professors i assignatures que incideixen sobre l'estudiant al llarg d'un període de temps<sup>32</sup>.
- Brow, Gary; Ehrman, Stephen Carhles; Lippincott, Joan K (2002) assenyalen que perquè l'avaluació sigui transformativa cal que: s'alineï amb els plans institucionals; promogui la col·laboració entre el personal d'administració i serveis, professorat, estudiants i comunitat; s'emprin fonts diverses i múltiples per demostrar els assoliments de l'estudiant (més enllà dels objectius específics de l'assignatura); que els resultats s'emprin sistemàticament per informar i remodelar l'ensenyament, i, finalment que es disseminin interna i externament aquests resultats.
- Alguns autors, com Angelo (1997) creuen que cal que els ensenyaments es converteixin en organitzacions que aprenen<sup>33</sup>. És a dir, en un entorn on tots els professors treballin junts, i reflexionin contínuament sobre allò que fan<sup>34</sup>. Cal que el professorat tingui una visió motivadora: de les fàbriques d'ensenyament a les comunitats d'aprenentatge, que canviï el seu model mental sobre el que és l'avaluació i la fonamenti en el coneixement que és l'aprenentatge profund, i que cada centre desenvolupi a partir dels seus propis actius les pròpies guies perquè l'avaluació promogui l'aprenentatge.

---

<sup>32</sup> Thomas Angelo (1997) assenyalava que, de fet, un dels motius pel quals l'avaluació no ha conduït a una major millora en els processos d'aprenentatge és a causa que les iniciatives han estat massa aïllades dins i entre les institucions, a banda de la manca de comprensió de què és l'aprenentatge i com es pot fomentar

<sup>33</sup> Aquest concepte és introduït per Peter Senge l'any 1992, en un llibre on descriu els principis que permetrien a les organitzacions *aprofitar l'entusiasme i la capacitat d'aprenentatge de la gent en tots els nivells de l'organització* (p12). El mateix autor, en una entrevista (O'Neil, 1995) afirmava que els ensenyaments estan molt allunyats d'aquest concepte, ja que els professors no treballen conjuntament sota una visió compartida d'allò que pretenen i no reflexionen sistemàticament sobre la seva pràctica diària.

<sup>34</sup> Precisament d'aquesta manera es defineix l'Scholarship (o activitat acadèmica) des de la Carnegie Foundation (Glassik, Huber i Maeroff, 1997), recollint el concepte de professional reflexiu (Schön, 1988). En aquest sentit, des de la investigació qualitativa hi ha metodologies com la investigació acció que no tenen altre objectiu que el d'involucrar el professorat en dinàmiques de reflexió sistemàtica amb l'objectiu de comprendre i millorar la seva pràctica educativa (Bartolomé 1995).