

Condicions per a una avaluació continuada que afavoreixi l'aprenentatge

Graham Gibbs
Claire Simpson



UNIVERSITAT DE BARCELONA



Quaderns de docència universitària 13

Títol original: *Conditions under wich assessment support students' learning*

Traducció al català: Francesc Jubany i Marta Fernández-Villanueva Jané

Director: Salvador Carrasco Calvo

Consell de redacció: Amelia Díaz Álvarez, Marta Fernández-Villanueva Jané, Gemma Fonrodona Baldajos, Mercedes Gracenea Zagarramundi, Evangelina González Fernández i Miquel Martínez Martín

Primera edició: març de 2009

© Graham Gibbs i Claire Simpson

© ICE i Editorial OCTAEDRO

Ediciones OCTAEDRO

Bailén, 5 - 08010 Barcelona

Tel.: 93 246 40 02 - Fax: 93 231 18 68

www.octaedro.com - octaedro@octaedro.com

Universitat de Barcelona

Institut de Ciències de l'Educació

Campus Mundet - 08035 Barcelona

Tel.: 93 403 51 75 – Fax: 93 402 10 61

Aquesta obra està sota llicència Reconeixement-NoComercial-SenseObraDerivada de Creative Commons. Podeu introduir, distribuir o comunicar públicament l'obra només sota els termes d'aquesta llicència. En cada còpia que reproduïu, distribuïu o comuniqueu públicament, hi heu de fer constar l'autor i la institució (ICE de la UB).

No podeu fer-ne ús comercial, ni tampoc obres derivades.

Et text complet de la llicència el podeu trobar a:

<http://creativecommons.org/licenses/by-nc-nd/2.5/es>

ISSN: 1988-1894

Disseny i producció: Serveis Gràfics Octaedro

ÍNDEX

Pròleg.....	5
Resum.....	9
Introducció.....	10
La influència determinant de l'avaluació.....	11
La importància de les activitats i els treballs del curs.....	14
El declivi de l'avaluació formativa.....	17
L'efectivitat del <i>feedback</i>	19
Com influir a través de l'avaluació sobre el volum, la selecció i la qualitat del que s'estudia.....	23
Condicció 1. Oferir als estudiants prou activitats avaluable per ocupar un temps d'estudi necessari i suficient.....	23
Condicció 2. Ocupar els estudiants amb activitats avaluable que els orientin sobre com distribuir correctament el temps i l'esforç en funció dels aspectes més importants de l'assignatura.....	25
Condicció 3. Afrontar una tasca avaluada implica assumir unes activitats d'aprenentatge productives adients.....	26
La influència del <i>feedback</i> en l'aprenentatge.....	29
Condicció 4. Un <i>feedback</i> suficient, tant pel que fa a la freqüència com al nivell de detall.....	29
Condicció 5. Orientar el <i>feedback</i> vers l'actuació dels estudiants, el seu aprenentatge i les accions que poden controlar, en comptes de centrar-lo en els estudiants o les seves característiques.....	31
Condicció 6. Oferir el <i>feedback</i> a temps per tal que els estudiants el rebin quan encara els importa i quan tenen temps per fer-lo servir de cara a un aprenentatge posterior o per rebre ajuda.....	31
Condicció 7. Un <i>feedback</i> adequat en relació amb la finalitat de l'activitat i amb els criteris d'avaluació.....	33
Condicció 8. El <i>feedback</i> és adequat en relació amb el que els estudiants creuen que han de fer.....	35
Concepcions de la tasca que tenen els estudiants.....	35
Concepcions de l'aprenentatge dels estudiants.....	35
Concepcions del coneixement que tenen els estudiants.....	36
La concepció del llenguatge de la disciplina en l'estudiant.....	37
Condicció 9. Recepció i anàlisi del <i>feedback</i>	38
Condicció 10. Resposta de l'estudiant vers el <i>feedback</i>	39

Conclusió.....	41
Agraïments.....	42
Bibliografia.....	43

Em comentava un estudiant d'una universitat politècnica espanyola que havia fet un curs acadèmic en una universitat nord-americana de prestigi, que la seva vida va fer un canvi radical a la universitat d'acollida i que va canviar qualitativament una altra vegada quan va tornar al nostre país. Entre els canvis que va viure allà destacava els següents: més treball personal amb menys assistència presencial a classe; menys continguts per aprendre; més temps per viure i passar-s'ho bé com a estudiant (més benestar) i un sistema d'avaluació diferent (fluid i eficaç), que consistia a oferir informació sobre el seu procés d'aprenentatge de manera continuada. A tots ens agrada que ens donin només la informació estratègicament necessària per entendre i comprendre allò de què es parla, que ens facin saber de manera ràpida i clara com aprenem i disposar de temps per viure i passar-nos-ho bé, a més de treballar i estudiar. Tanmateix, l'estudiant continuava subratllant la importància d'un sistema d'avaluació que facilita contínuament informació sobre la manera com apren l'estudiant i d'una bona selecció de continguts, escollits segons el valor estratègic i no segons el volum que ocupen, com a factors clau per al benestar de l'estudiant.

Al meu parer, aquests factors no només són clau per al benestar de l'estudiant que vol aprendre, sinó també per optimitzar la qualitat de la docència a la universitat. I no crec que siguin conseqüència ni de la implantació de l'Espai Europeu d'Educació Superior ni de cap dels documents derivats de la Declaració de Bolonya. Senzillament són conseqüències lògiques de qualsevol projecte de millora de la qualitat de la docència a la universitat. La millora de la qualitat de la docència no s'assoleix només fixant uns bons objectius i regint-nos per resultats d'aprenentatge i competències, sinó que exigeix sobretot rigor a l'hora d'escollir els continguts en funció del valor estratègic i cura en l'avaluació. Convé aprofitar la integració al procés de convergència europea per replantejar la nostra feina com a docents i dedicar temps a aquestes qüestions. Un cop ajustats i, en el millor dels casos, construïts els plans docents dels nous graus i postgraus, cal dedicar temps acadèmic a escollir els millors continguts i a posar en pràctica bons sistemes d'avaluació continuada.

Els canvis que acompanyen el procés de construcció de l'Espai Europeu d'Educació Superior poden aprofitar-se per generar condicions que fomentin l'aprenentatge amb profunditat en els contextos de docència universitària. Entre aquests canvis, n'hi ha que són motivats per la mateixa dinàmica del procés d'elaboració dels nous títols de grau i postgrau —sempre que aquesta tasca no s'hagi limitat a una reestructuració purament superficial en funció d'un ajustament dels resultats en termes de recursos docents per departament—, com ara la pràctica de la planificació docent per matèries i la programació per assignatures. N'hi ha d'altres que són conseqüència de la integració de les tecnologies de la informació i la documentació al servei de l'aprenentatge i la comunicació i que permeten avançar vers sistemes d'aprenentatge i docència a la universitat de tipus híbrid, que combinen situacions presencials amb d'altres que no ho són. També n'hi ha d'altres que estan lligats a la noció de «bon nivell d'aprenentatge» en una societat com ara l'actual, caracteritzada per un creixement exponencial de la informació als diferents àmbits del coneixement, fins a l'extrem que cal escollir, entre tots els continguts al nostre abast susceptibles de ser ensenyats, aquells que generen un millor aprenentatge. Ara bé, per assolir un aprenentatge millor i amb profunditat cal que la planificació docent, la programació, la integració de tecnologies per a l'aprenentatge i la comunicació i la selecció de continguts tinguin al costat sistemes d'avaluació que fomentin un aprenentatge fructífer a llarg termini i equips docents disposats a dur a terme aquesta feina.

Es pot afirmar, perquè hi ha evidències que ho corroboren, que les assignatures en què els estudiants obtenen més bons resultats són aquelles en què hi ha una estreta col·laboració entre el professorat que les imparteix. En canvi, les assignatures amb unes qualificacions més baixes són aquelles en què el professor s'ocupa únicament de la seva assignatura. Allò que determina el resultat global de l'aprenentatge no és la suma d'espais docents estancs —per molt bona que sigui la docència en cadascun d'aquests espais—, sinó de la feina conjunta del professorat en el marc del pla docent de la titulació. Per això, el treball en equips docents és clau. Els equips docents permeten discutir i reflexionar sobre la pràctica del professorat que els integren, una transferència més àgil de la innovació que contribueix a la millora de la qualitat i facilita l'avaluació conjunta.

No obstant això, el professorat està acostumat a fer la seva feina docent i avaluadora en clau individual i en el context de l'assignatura. El canvi vers una cultura docent que integri la feina del professorat mitjançant equips docents no és fàcil, però és imprescindible. Per això, cal generar espais de formació i anàlisi en què el professorat que hi participa pugui reflexionar sobre la seva pràctica docent i construir coneixement sobre els processos d'aprenentatge, docència i avaluació a la universitat de manera col·laborativa. Espais de formació o intercanvi entre professorat de disciplines afins o de disciplines diferents que afavoreixin capacitats i actituds de cooperació entre els participants i desenvolupin les seves competències a l'hora d'escollir els millors continguts, comunicar millor el coneixement i conèixer les funcions estratègiques de l'avaluació que permetin als docents escollir les millors tàctiques d'avaluació.

Aquesta darrera qüestió, l'avaluació dels aprenentatges dels estudiants, és la que tracta el text que presentem. Des de l'Institut de Ciències de l'Educació, i mitjançant la nostra col·lecció *Quaderns de docència universitària*, ens proposem amb aquest quadern oferir un conjunt de reflexions i una revisió sobre les millors condicions per aplicar una avaluació continuada que afavoreixi l'aprenentatge dels estudiants.

La feina de Graham Gibbs i Claire Simpson que ara presentem planteja una revisió excel·lent dels estudis sobre l'avaluació dels aprenentatges a la universitat. El text proposa una sèrie de condicions en què l'avaluació afavoreix l'aprenentatge i les argumenta amb referències teòriques, proves empíriques i experiències pràctiques. Marta Fernández Villanueva, professora de la Facultat de Filologia de la Universitat de Barcelona, i fins fa uns mesos coordinadora del programa d'assessorament en formació per a la innovació docent de l'ICE de la UB, ha traduït el text i ha adaptat al nostre llenguatge acadèmic i context universitari les consideracions i reflexions que hi apareixen. Volem agrair a en Graham i la Claire el seu estudi i disponibilitat i a la Marta, la seva bona feina.

Teniu a les mans un text àgil que, com apunten els seus autors, no parla dels sistemes que utilitzem per mesurar els resultats, sinó de l'aprenentatge. Un text que defensa i argumenta que el més important de

l'avaluació és que afavoreixi un aprenentatge fructífer i que de la fiabilitat cal ocupar-se'n després. Al text se subratlla la influència cabdal de l'avaluació sobre l'aprenentatge dels estudiants, es constata el retrocés de l'avaluació formativa en algunes universitats a causa de la manca de recursos i es proposen i argumenten deu condicions per a una avaluació continuada de l'aprenentatge.

Penso que és un llibre fantàstic per fomentar l'anàlisi, la reflexió i el debat entre el professorat sobre el valor estratègic de l'avaluació, les concepcions que tenen els estudiants sobre l'aprenentatge i l'avaluació i per avançar vers tàctiques i pràctiques d'avaluació de la matèria compartides per l'equip docent.

Universitat de Barcelona, gener de 2009.

Miquel Martínez

Director de l'Institut de Ciències de l'Educació

RESUM

Una bona part de l'avaluació de la docència se centra en la tasca que els docents fan a classe. Aquest article, en canvi, s'ocupa d'avaluar la manera com es fa l'avaluació i com influeix en la manera d'aprendre i d'estudiar fora de classe. En general, es parteix de la base que la manera d'avaluar condiciona especialment el contingut, la manera i la intensitat d'estudi dels alumnes. Aquest article proposa un conjunt de «condicions amb les quals l'avaluació pot afavorir l'aprenentatge» i les argumenta amb referències teòriques, proves empíriques i experiències pràctiques. A més a més, aquestes condicions ofereixen un marc de referència als docents disposats a revisar l'efectivitat de la seva pròpia pràctica avaluadora.

INTRODUCCIÓ

Quan la docència universitària apareix als titulars dels diaris és gairebé sempre per qüestions relacionades amb l'avaluació: són notícia una suposada baixada del nivell, la detecció de casos de plagi, errors als exàmens, etc. L'actuació més recent de l'Agència de la Qualitat de l'Educació Superior del Regne Unit (QAA) s'ha centrat en els resultats de l'aprenentatge i en la manera d'avaluar-los, a més d'especificar els nivells estàndard i el paper dels examinadors per garantir-los, mentre que les mesures adoptades com a estratègies institucionals d'educació i aprenentatge aplicades a l'avaluació prioritzen, pràcticament sempre, el fet de casar els resultats de l'aprenentatge amb l'avaluació i d'especificar els criteris d'aquesta avaluació. El centre d'interès de tots els agents, des dels mitjans de comunicació fins a les institucions, passant pels responsables de qualitat, és l'avaluació entesa com un sistema per mesurar resultats. En aquest article, però, no ens interessen aquests sistemes, sinó l'aprenentatge. Els sistemes d'avaluació més segurs, fiables i rigorosos són sovint companys de viatge d'aprenentatges superficials i estèrils i estan vinculats a resultats d'aprenentatge de curta durada (de fet, moltes vegades són la via directa a aquest tipus d'aprenentatge superficial). Amb això no volem defensar una avaluació poc fiable, però estem convençuts que els sistemes d'avaluació s'han de dissenyar pensant bàsicament a fomentar un aprenentatge fructífer a llarg termini; de la fiabilitat cal ocupar-se'n després. Els nivells pujaran si millora l'aprenentatge dels estudiants i no si millora la manera de mesurar un aprenentatge limitat. Per això, aquest article aborda el disseny d'una avaluació que fomenti l'aprenentatge amb profunditat. Els estudis de cas publicats sobre aquest tema descriuen mètodes d'avaluació concrets, les tàctiques emprades, mentre que en d'altres publicacions podem trobar orientacions sobre la manera de dur a la pràctica un ampli ventall d'aquestes tàctiques (per exemple, Gibbs, 1995). Aquest article analitza l'estratègia, les funcions de l'avaluació que permeten als docents escollir les tàctiques d'avaluació més escaients. Demostrarem que l'avaluació és més efectiva per a l'aprenentatge si es compleixen una sèrie de condicions. En aquest article, ens interessarem per aquestes condicions.

LA INFLUÈNCIA DETERMINANT DE L'AVALUACIÓ

A començament dels anys setanta, estudiosos de tots dos costats de l'Atlàntic (Snyder, 1971; Miller i Parlett, 1974) es van dedicar a investigar com era l'aprenentatge dels estudiants en diferents universitats de prestigi. Amb aquesta recerca van descobrir que, sorprenentment, el que més incidia en els estudiants no era la docència sinó l'avaluació. Els estudiants descrivien tots els aspectes de la seva activitat com a estudiants (les classes a què anaven, el volum de feina que feien i com els anava als estudis) com si estiguessin totalment determinats per la manera com percebien les exigències del sistema d'avaluació. Derek Rowntree va arribar a firmar que «si volem descobrir la veritat sobre el sistema educatiu, primer hem d'estudiar els procediments d'avaluació» (Rowntree, 1987, p. 1). Els estudis de Snyder i Miller, i Parlett van ser els que van anar més lluny i van revelar com responen els estudiants a aquests procediments d'avaluació. Més recentment, diversos estudis qualitius han fet èmfasi en la importància d'entendre com responen els estudiants a les innovacions en l'avaluació (Sambell i McDowell, 1998).

La tasca de Snyder va establir les bases de la noció de «currículum ocult», en contrast amb el currículum formal presentat a la documentació de l'assignatura, que els estudiants han d'anar descobrint i observant si se'n volen sortir:

Des del començament, vaig tenir la impressió que tot consistia en una mena d'exercici d'administració del temps... Cal extreure el que és realment important de cada assignatura... Materialment, no es pot fer tot. Em vaig adonar que, si aconsegueixes aïllar el més important, llavors pots aspirar a tirar endavant totes les assignatures sense cap problema. (Snyder, 1971, p. 62-63)

Un cop esbrinat en què consisteix aquest currículum ocult, els estudiants poden distribuir els esforços d'una manera més eficient:

Ara no em preocupo per fer els exercicis, sinó que plantejo les assignatures de manera que pugui treure un excel·lent de la manera més fàcil i

sembla mentida la poca feina que cal fer si realment no t'agrada l'assignatura. (Snyder, *ib.* p. 50)

Miller i Parlett es van centrar en analitzar fins a quin punt els estudiants actuaven en funció dels indicis vinculats a les recompenses que ofería el sistema d'avaluació. Així, van identificar diferents tipus d'estudiants, com ara el «buscador d'indis», especialitzat en fer «confessar» al docent el contingut de l'examen o les seves preferències personals, el «conscient dels indis», que escoltava i parava atenció als consells que li oferien els docents sobre què era important i què no ho era, i el «sord als indis», que desestimava sistemàticament totes les orientacions rebudes. Un d'aquests estudiants «buscadors d'indis» descriu el procés per descobrir les preguntes de l'examen:

Estic convençut que es tracta d'un joc, del joc d'examinar-se. Algunes dades, per exemple, no les estudies, i tampoc no vas a totes les classes. Hi vas, mires els exàmens i penses: «Sembla que hi havia quatre preguntes sobre un tema concret aquest any, i l'any passat el professor va dir que l'examen seria més o menys semblant al de l'any anterior». Així, de seguida pots descartar una bona part de l'assignatura... (Miller i Parlett, 1974, p. 60)

En canvi, els estudiants batejats com a «sords als indis» van manifestar l'opinió següent:

No selecciono els temes a l'hora de preparar l'examen. No em sento segur si només em limito a uns determinats temes.
Miro d'estudiar-m'ho tot. (Miller i Parlett, 1974, p. 63)

Miller i Parlett van arribar a predir amb una gran exactitud quins estudiants obtindrien millors resultats en la graduació:

Les persones «conscients dels indis» tenien tendència a arribar a ser *Upper Seconds*, mentre que els «sords als indis» se situaven com a *Lower Seconds*.¹ (Miller i Parlett, 1974, p. 55)

1. A les universitats del Regne Unit, en funció dels resultats obtinguts, la qualificació es considera de primera classe (+70%), segona superior (60-69%), coneguda també com a «2:1» o *Upper Second Class*, segona inferior (50-59%), o «2:2» o *Lower Second Class*, tercera classe (40-49%), «3» o *Third Class*, insuficient (30-39%) o molt deficient (0-29%).

Molts estudiants són ben capaços de diferenciar entre els punts als quals han de dedicar més atenció per motius d'avaluació i els punts que són més efectius per a un aprenentatge profund, tal com explica aquest estudiant d'oceanografia:

Si estàs sotmès a molta pressió, llavors et concentres en aprovar l'assignatura. Ho sé per experiència, per una mala experiència. En una matèria que no dominava prou vaig intentar entendre els temes i vaig acabar suspentent l'examen. Quan vaig tornar a fer l'examen, simplement em vaig centrar en superar l'examen i prou i vaig treure un 96%; el professor no entenia com havia pogut suspendre la primera vegada. Jo li vaig dir que m'havia centrat només en passar l'examen i no pas en entendre la matèria. Continuo sense entendre la matèria, és a dir que d'alguna manera no he aconseguit el que volia. (Gibbs, 1992, p. 101)

Malgrat que la documentació de l'assignatura especifiqui clarament què cal avaluar, els estudiants descobreixen amb els seus propis mitjans què compta més (o si més no què creuen que compta més) i orienten els seus esforços en aquesta línia. Per tant, són estratègics pel que fa a l'ús del temps i «selectivament negligents» en la mesura que eviten continguts que es pensen que no s'avaluaran. Hi ha qui considera que els estudiants, des dels anys setanta, s'han tornat més estratègics pel que fa a l'ús del temps i les energies i que cada vegada estan més condicionats, i no pas menys, per les exigències del sistema d'avaluació, en el sentit que planifiquen més el seu pas pels estudis (MacFarlane, 1992).

LA IMPORTÀNCIA DE LES ACTIVITATS I ELS TREBALLS DEL CURS

Els estudiants tenen tendència a obtenir més bones qualificacions als treballs que no pas als exàmens (James i Fleming, 2004-2005). Chansarkar i Raut-Roy (1987) van estudiar els efectes de la combinació de diferents formes de treballs i activitats amb exàmens i van detectar que totes les combinacions de treballs i activitats amb exàmens es traduïen en més bones notes de mitjana que no pas en els casos en què només es feien exàmens (les qualificacions eren de mitjana un 12% més altes).

Gibbs i Lucas (1997) van analitzar les qualificacions de 1.712 mòduls de la Universitat Politècnica d'Oxford. Als mòduls dissenyats al 100% amb activitats i treballs, les qualificacions mitjaneres eren un 3,5% superiors a les qualificacions obtingudes als mòduls en què només es feien exàmens. Una de les conclusions observades va ser la correlació positiva entre la proporció d'activitats i treballs d'un mòdul i la mitjana de les qualificacions ($r = +0,36$, $p < 0,0001$). Bridges *et al.* (2002) van estudiar les diferències en les qualificacions obtingudes amb activitats i treballs i en exàmens en sis titulacions de quatre universitats i van detectar que les qualificacions obtingudes en activitats i treballs del curs eren més altes en una tercera part de la classificació per titulacions d'estudis anglesos i història (una conclusió semblant a l'observada a la Universitat Politècnica d'Oxford) i més alta en dues terceres parts en biologia, empresarials, enginyeria de sistemes i dret.

De fet, els estudiants prefereixen fer activitats i treballs, tal com indicava ja Starr (1970) quan afirmava que el 90% dels estudiants de quatre departaments preferia que la meitat o més de les qualificacions s'obtingués a partir d'activitats i treballs. A més a més, un 56% preferia que totes les qualificacions s'obtinguessin a partir de treballs. Els estudiants consideren que els treballs i les activitats de curs constitueixen una manera més justa de valorar la feina que no pas els exàmens, ja que valoren un ventall més ampli de competències que els exàmens i perquè els permeten organitzar-se millor la feina (Kniveton, 1996).

El fet que la mitjana de les qualificacions sigui més alta o que les preferències dels estudiants tendeixin cap a una determinada línia no serien arguments suficients si les activitats i els treballs de curs fossin inherentment menys vàlids com a forma d'avaluació acreditativa. Però no ho són. Per començar, els exàmens no ajuden massa a l'hora de predir possibles actuacions posteriors, com ara la inserció o l'èxit professional. Una anàlisi de 150 estudis centrats en la relació entre resultats d'exàmens i un ampli conjunt d'indicadors d'èxit en la vida adulta posa de manifest que la relació és, en el millor dels casos, poc clara (Baird, 1985). Per exemple, els resultats del primer cicle expliquen menys del 10% de la variació en l'actuació en els postgraus (Warren, 1971).

En segon lloc, les qualificacions de les activitats i els treballs de curs representen una eina més vàlida per predir l'aprenentatge profund i a llarg termini del contingut de les assignatures que no pas els exàmens. Conway *et al.* (1992) van estudiar els resultats dels estudiants de psicologia, a través de diferents proves per avaluar-ne la capacitat d'entendre i recordar el contingut d'una assignatura de psicologia cognitiva que havien fet molts anys enrere. Una de les observacions va ser que les qualificacions obtingudes en activitats i treballs de curs fets fins a 13 anys enrere mostraven una correlació amb els resultats d'aquestes proves, a diferència de les qualificacions dels exàmens. Podem deduir que el tipus d'aprenentatge que impliquen les activitats i els treballs de curs té efectes a llarg termini, mentre que el tipus d'aprenentatge de l'estudi per a un examen no en té. Els estudis sobre aprenentatge superficial o en profunditat posen de manifest uns resultats semblants: és a dir, que els efectes positius en els resultats dels exàmens dels estudiants que optaven per un aprenentatge superficial a l'hora de preparar els exàmens només es mantenien a curt termini (Marton i Wenestam, 1978).

En tercer lloc, estudis experimentals en els quals els estudiants no s'havien preparat ni per als exàmens ni per fer activitats i treballs de curs han demostrat que la qualitat de l'aprenentatge és més alta en les assignatures basades en activitats i treballs. Per exemple, Tynjala (1998) ha comparat dos grups d'estudiants: el primer estudiava amb classes expositives convencionals, a partir d'un manual i fent un examen, mentre que el segon grup estudiava a través d'activitats basades en el manual, amb debats amb els altres estudiants al voltant de les

activitats i elaborant un treball escrit que havia d'avaluar el professor. Aquest segon grup també va fer l'examen final per tal de poder comparar els resultats amb el primer grup, malgrat que no havien estudiat per a l'examen. Aleshores es va observar que el segon grup feia més èmfasi en l'elaboració conceptual i que havia desenvolupat concepcions més complexes del significat de l'aprenentatge (Säljö, 1982). A les respostes dels exàmens recorrien més a les comparacions, les valoracions i les construccions més complexes en termes de taxonomia, però només sobre resultats d'aprenentatge (Biggs i Collis, 1982). Aquests resultats, obtinguts amb menys hores de docència, es van interpretar com uns productes més ben construïts tenint en compte els requisits d'avaluació del segon grup.

Entre els docents universitaris, hi ha l'opinió generalitzada que si se suprimeixen activitats i treballs d'un mòdul a causa de restriccions de mitjans, els estudiants simplement deixen de fer l'estudi associat. Per exemple, els estudiants ben poques vegades faran un treball que no s'hagi d'avaluar. Un dels arguments és que cal avaluar-ho tot per aconseguir que els estudiants dediquin temps i esforç a dur-ho a terme. Tanmateix, les activitats i els treballs de curs no necessàriament s'han de qualificar per tal que es generi l'aprenentatge necessari. Forbes i Spence (1991) descriuen la recerca duta a terme en les pràctiques d'avaluació d'una assignatura d'enginyeria de la Universitat de Strathclyde. Quan els docents van deixar de puntuar els exercicis de problemes setmanals (simplement perquè estaven desbordats), els estudiants també van deixar de fer els problemes i, com a conseqüència d'això, les notes dels exàmens van baixar. En canvi, quan els docents van introduir l'avaluació paritària periòdica dels exercicis de problemes (com a activitat obligatòria del curs però sense contribuir a la nota final) els resultats dels exàmens van pujar espectacularment fins a un nivell molt superior al que havien assolit abans, quan els docents puntuaven els exercicis. El desllorigador de l'aprenentatge, en aquest cas, no va ser la gran quantitat de correccions, sinó la qualitat de la implicació dels estudiants en activitats d'aprenentatge. L'element clau a l'hora de dissenyar sistemes d'avaluació és assolir la implicació dels estudiants en activitats d'aprenentatge sense necessitat de generar muntanyes de paper per corregir.

EL DECLIVI DE L'AVALUACIÓ FORMATIVA

Un dels trets característics tradicionals de la docència universitària al Regne Unit ha estat la capacitat d'aportar *feedback* personalitzat i detallat sobre les activitats i els treballs de curs d'una manera regular. El cas paradigmàtic és el de les universitats d'Oxford i Cambridge, en què els estudiants redactaven un treball cada setmana i el llegien al seu tutor durant la tutoria individual, en la qual rebien, de manera immediata i detallada, els comentaris crítics relacionats amb l'elaboració conceptual presentada. Aquest tipus d'aprenentatge era tot el que rebien molts estudiants d'aquestes institucions: ensenyar equivalia a aportar *feedback* sobre els treballs. Es tractava d'una avaluació formativa bastant separada de l'avaluació sumatòria i, a Oxford i Cambridge, l'única avaluació sumatòria es limitava moltes vegades als exàmens finals després de tres anys d'estudi, un període en què s'havia treballat a partir de l'avaluació formativa setmanal.

Ben poques institucions han aconseguit arribar a la quantitat o la qualitat del *feedback* que ofereixen les universitats d'Oxford i Cambridge, tot i que en general s'acceptava que les activitats i treballs freqüents, juntament amb un *feedback* detallat (per escrit), eren fonamentals per a l'aprenentatge dels estudiants. Fins fa relativament poc temps, per exemple, moltes titulacions de ciències integraven també exercicis setmanals de problemes i informes de laboratori, en tots els casos avaluats i puntuats per docents i retornats als estudiants en el termini d'una setmana. En la major part de modalitats d'educació a distància, el *feedback* sobre les activitats més habituals és el component interactiu més important de la docència i l'Open University ha fet un gran èmfasi en les activitats freqüents, en la formació i la remuneració de tutors per oferir un *feedback* complet i detallat i en un control i un seguiment que garanteixin la qualitat d'aquest *feedback*. Per a alguns estudiants de l'Open University, aquest és l'únic contacte que tenen amb el seu tutor: poden passar sense massa contacte presencial, o fins i tot sense gens ni mica, però no poden passar sense un *feedback* periòdic sobre les activitats i les tasques fetes.

A les universitats convencionals, les limitacions de recursos s'han traduït en una reducció de la freqüència de les activitats i les tasques, com

també de la qualitat i la quantitat de *feedback* ofert, i en un augment del temps de resposta més adequat per oferir *feedback*. La modularització ha tingut tendència a escurçar el temps de les assignatures i ha reduït l'escala temporal que permetia proposar activitats i oferir *feedback*, a banda d'augmentar el nombre d'exàmens. En alguns cursos s'han abandonat totalment les activitats pensades per a l'avaluació formativa, mentre que en altres potser es manté una única activitat o tasca, però amb un *feedback* molt dilatat en el temps, de vegades posterior fins i tot als exàmens. Simultàniament, la diversitat d'estudiants ha crescut notablement, de manera que es fa difícil parlar amb certesa sobre el nivell de complexitat dels coneixements previs, les tècniques i els hàbits d'estudi, la concepció mateixa de l'aprenentatge (Säljö, 1982) o la concepció del coneixement (Perry, 1970) que aporten els estudiants. Semblaria més aviat que cal potenciar l'orientació i oferir pautes als estudiants que necessiten més experiència per dur a terme activitats i tasques, i més *feedback* sobre l'aprenentatge, no pas menys. Pel que fa a l'educació a distància, en canvi, i precisament perquè les activitats periòdiques i el *feedback* detallat s'hi consideren elements fonamentals, s'han mantingut àmpliament i, com a conseqüència d'això, els estudiants de l'Open University és possible que avui rebin 50 vegades més *feedback* sobre les seves activitats d'aprenentatge durant els estudis que els estudiants de les universitats convencionals.

L'EFFECTIVITAT DEL *FEEDBACK*

En el marc d'una revisió exhaustiva de 87 metaanàlisis de projectes de recerca sobre els elements més significatius a l'hora d'obtenir bons resultats d'aprenentatge, Hattie (1987) va detectar que el factor individual més potent era el *feedback*. La revisió dels treballs sobre avaluació formativa que ofereixen Black i William (1998) destaca d'una manera semblant els efectes positius sobre l'aprenentatge, tangibles i de llarg abast, que el *feedback* ofereix en comparació amb altres aspectes de la docència. S'han produït molts intents per comprendre tant la naturalesa d'aquest impacte com la possibilitat d'aprofitar-ne el potencial a través de la innovació, si més no en les escoles, arran d'aquesta prova irrefutable.

A l'educació superior, el *feedback* individual que s'ofereix als estudiants a classe mostra una clara tendència a la baixa a mesura que augmenten les dimensions del grup, tot i que aquest fet no s'ha documentat. Els comentaris escrits sobre les activitats i les tasques, tanmateix, continuen figurant entre els principals components de la dedicació del professorat a l'educació superior. Amb l'augment de la mida dels grups s'han anat introduint mesures d'economia d'escala en la docència (simplement augmentant el nombre d'estudiants per aula), però les economies d'escala són difícils de materialitzar amb èxit en l'avaluació: els esforços de l'avaluació augmenten en proporció directa al nombre d'estudiants. Com a conseqüència d'això, els esforços d'avaluació poden acabar acumulant la major part dels esforços de la docència i els docents poden acabar dedicant la major part del temps a corregir i posar notes. Val la pena tot aquest esforç?

En el mar del Course Experience Questionnaire (Ramsden, 1991), molt emprat a Austràlia i altres països per avaluar la qualitat dels cursos, el punt del qüestionari que diferencia clarament entre els millors cursos i els pitjors és el que afirma que «el personal docent aquí ofereix *feedback* útil sobre l'aprenentatge» (Ramsden, 1992, p. 107). Això no vol dir que els docents universitaris ofereixin realment *feedback* útil, sinó només que el fet d'oferir-ne és el més significatiu de la seva aportació. Fins a quin punt en realitat és efectiu el fet d'oferir *feedback*?

MacLellen (2001) va preguntar a 130 estudiants i 80 docents de la Universitat de Strathclyde sobre la seva percepció de l'avaluació. De les 40 preguntes plantejades, quatre tenien relació amb el *feedback* i van posar de manifest unes importants discrepàncies entre estudiants i docents. Mentre la major part dels docents responia que el *feedback* detallat és útil «sovint», ajuda a l'estudiant a entendre les coses «sovint» i contribueix a l'aprenentatge «sovint», la majoria dels estudiants va respondre que el *feedback* només era útil «de vegades» en aquests casos. El 30% dels estudiants va respondre que el *feedback* «mai» no els ajuda a entendre les coses. Així mateix, si el 63% dels docents responia que el *feedback* convida «sovint» a consultar el tutor, només el 2% dels estudiants responia en el mateix sentit, mentre que el 50% responia que el *feedback* no convida «mai» a consultar amb el tutor.

És possible que tinguem un problema tan important amb la quantitat i la qualitat del *feedback* que, efectivament, l'efectivitat del *feedback* que reben els estudiants quedi invalidada. Al cap i a la fi, els docents treballen sota una enorme pressió de temps i de feina i, en aquestes circumstàncies, es fa difícil oferir *feedback* en profunditat, detallat i útil. Tanmateix, també hi ha altres problemes. Una lectura dels estudis sobre què fan els estudiants amb el *feedback* pot provocar autèntica depressió. Tot sovint, ni tan sols el llegeixen (Hounsell, 1987) o no l'entenen (Lea i Street, 1998). Wotjas (1998) indica:

Alguns estudiants, si no estan d'acord amb la nota, llencen el *feedback* a la paperera, mentre que d'altres només semblen interessats en el resultat final i ni tan sols passen a recollir la feina corregida.

A més a més, hi ha un altre problema, relacionat tant amb les notes com amb el *feedback* que s'ofereix. L'estudiant possiblement percep que la qualificació és un indicador de la seva habilitat personal, o fins i tot una valoració de la seva persona, perquè en general es tracta d'una avaluació normativa que indica on se situa cadascú en relació amb els altres. Una qualificació baixa pot malmetre «l'autoeficàcia» d'un estudiant, és a dir, la sensació que té en relació amb la seva pròpia eficàcia. Yorke (2001) abunda en les maneres positives i negatives com l'avaluació formativa pot afectar el manteniment dels estudiants i fa èmfasi en el paper que pot tenir en la «integració acadèmica» (Tinto, 1993). En

canvi, el *feedback* és més possible que es percebi per part seva com un comentari sobre el que s'ha après. Quan no hi ha qualificacions, s'ha demostrat que els estudiants llegeixen el *feedback* amb molta més atenció (Black i William, 1998) i el fan servir per guiar l'aprenentatge. Tal com demostren els estudis (en l'àmbit escolar), alguns centres han adoptat la política d'oferir només *feedback* sobre les activitats o les feines fetes, sense puntuar-les. El sistema conegut com a «avaluació com a aprenentatge», de l'Alverno College, probablement és l'exemple més conegut, en el camp de l'educació superior, d'avaluació «amb menys notes».

El paisatge no és massa prometedor. D'una manera o d'una altra, l'avaluació exigeix un esforç enorme, desagrada tant a estudiants com a professors i es mostra molt ineficaç a l'hora de plantejar i justificar un seguit de «condicions en què l'avaluació pot afavorir l'aprenentatge». Hi ha pocs casos en què les proves siguin definitives com per argumentar que si l'avaluació compleix unes determinades condicions, aleshores l'aprenentatge per força serà més efectiu. Per tant, es plantegen més aviat com un conjunt d'orientacions plausibles.

No ens trobem pas davant del primer intent d'identificar unes «condicions» d'aquesta mena, tot i que sí que ho seria en el context de l'educació superior. Els estudis en l'àmbit escolar ja han identificat els efectes de l'avaluació formativa i els han sistematitzat tal com es pot observar a la llista següent, basada en Gagne (1977):

1. Reactivar o consolidar competències o coneixements previs necessaris abans d'introduir noves matèries.
2. Centrar l'atenció en els aspectes més importants de la matèria.
3. Incentivar les estratègies d'aprenentatge actiu.
4. Oferir als estudiants oportunitats per posar en pràctica les seves competències i consolidar l'aprenentatge.
5. Difondre els resultats i oferir *feedback* de correcció.
6. Ajudar els estudiants a controlar la seva evolució i desenvolupar competències d'autoavaluació.
7. Orientar-los en la presa de decisions sobre les properes activitats de formació o aprenentatge per millorar el domini.
8. Ajudar els estudiants a percebre que han assolit un objectiu.

Les condicions plantejades exerceixen dues menes d'influències, de caire lleugerament diferent:

- La influència, a través del disseny dels sistemes d'avaluació i de les activitats i els treballs, sobre com estudien, què estudien i de quina manera s'impliquen els estudiants.
- La influència, a través del *feedback*, sobre l'aprenentatge.

COM INFLUIR A TRAVÉS DE L'AVALUACIÓ SOBRE EL VOLUM, LA SELECCIÓ I LA QUALITAT DEL QUE S'ESTUDIA

Condicció i

Oferir als estudiants prou activitats avaluable per ocupar un temps d'estudi necessari i suficient

Aquest tema està relacionat amb el temps i l'esforç que dediquen els estudiants a l'objectiu d'aprenentatge (el principi de «temps dedicat a una tasca», present a Chickering i Gamson, 1987), segons el qual si els estudiants no dediquen prou temps a una cosa, no l'arribaran a aprendre. Berliner (1984), en el resum de l'estudi dut a terme sobre el «temps dedicat a una tasca», va arribar a la conclusió que hi ha proves empíriques més que suficients per establir una relació directa entre temps dedicat a les assignatures, gestió del temps de l'estudiant i temps real de l'estudiant dedicat a la realització de la tasca, d'una banda, i els resultats de l'aprenentatge de l'estudiant, de l'altra.

La relació entre esforç i qualificació no sempre és directa. Kember *et al.* (1996) van detectar que les percepcions que tenien els estudiants sobre el seu esforç depenien més de la seva motivació que del nombre d'hores que realment hi dedicaven i que hi havia estudiants que «feien moltes hores» improductives, especialment si adoptaven un estil d'aprenentatge superficial a l'hora d'estudiar. Determinats tipus d'avaluació poden generar moltes hores de memorització ineficaç.

El disseny de les assignatures a les universitats del Regne Unit té en compte un nombre determinat d'hores d'aprenentatge relacionat amb el nombre de crèdits de l'assignatura. Normalment, els estudiants han de dedicar a l'estudi entre una i quatre hores per cada hora passada a classe (en funció, en gran mesura, de la disciplina en qüestió). Innis (1996) va observar que els estudiants de la Leeds Metropolitan University dedicaven entre 1,4 i 3,0 hores d'estudi per cada hora de classe. La part d'aquest temps «fora de classe» que es dedica realment a estudiar està molt condicionada per les exigències de la forma d'avaluació. Als Estats Units, els estudiants universitaris dediquen de mitjana a l'estu-

di menys de la meitat del temps que n'esperen els seus professors: entre 0,3 i 1,0 hores d'estudi fora de classe, mentre que els docents n'esperen, de mitjana, 2,1 hores per cada hora de classe (Moffat, 1989; Hutchings *et al.*, 1991; Gardiner, 1997; Brittingham, 1988). Als Estats Units, sempre que s'ha provat de millorar els resultats dels estudiants a través de l'avaluació, s'ha fet èmfasi en l'anomenada «avaluació a classe» (és a dir, activitats fetes a classe per avaluar els estudiants, normalment a través de qüestionaris) i aquesta informació aportada per l'avaluació s'empra com a eina d'orientació tant per a estudiants com per a docents (Angelo i Cross, 1993). Aquesta atenció especial per les activitats de la classe es pot interpretar com un reconeixement del fracàs a l'hora de generar gran part de l'aprenentatge fora de classe, a través del tipus d'avaluació que fan servir. Els estudis fets a través de diaris d'aprenentatge (com ara Innis, 1996) demostren que els estudiants del Regne Unit dediquen principalment el seu temps a les activitats que s'avaluaran i que això, amb el temps, es tradueix en un focus d'atenció cada vegada més limitat, a mesura que van adquirint experiència com a estudiants, fins que arriben a l'extrem de dedicar, pels volts del tercer any, només un 5% del temps a activitats d'estudi no relacionades amb l'avaluació.

Les àrees disciplinàries amb activitats avaluades amb menys freqüència (com ara les que es basen en textos) són les que tenen els estudiants que hi dediquen menys hores (Vos, 1991). Les àrees de ciències i tecnologia, que impliquen una major dedicació a l'estudi, normalment programen activitats que s'avaluen amb més freqüència (i que també són més limitades), com ara fulls d'exercicis de problemes i informes de laboratori.

Els estudis duts a terme per esbrinar com influeix la feina remunerada exercida en paral·lel a la dedicació a l'estudi a temps complet demostren que aquests estudiants dediquen menys hores (Curtis i Shami, 2002) i que els seus resultats són significativament pitjors (Paton-Salzberg i Lindsay, 1993). Aquests estudis plantegen que més de tres quartes parts dels estudiants a temps complet treballen durant el període lectiu i que distribueixen les seves hores d'estudi reduïdes d'una manera estratègica en relació amb els requisits de l'avaluació. Així mateix, indiquen que redueixen les lectures i altres activitats d'estudi fora de classe.

Proposar activitats i treballs que s'avaluin no és l'única manera de captar el temps i l'esforç dels estudiants. La manera convencional de fer-ho és a través d'exàmens amb una selecció imprevisible de continguts de l'assignatura i sense avisar, de manera que per a un estudiant el fet de no estar al dia representi una pràctica d'alt risc. La qualitat, la quantitat i la distribució de l'esforç d'estudi que s'aconsegueixen d'aquesta manera són imprevisibles i probablement varien en funció de la percepció que els estudiants tenen de les demandes que els pot plantejar l'examen i dels riscos que implica.

També es pot aspirar que els estudiants dediquin el temps i l'esforç necessari a través de la pressió social, apel·lant per exemple a:

- El possible ridícul davant els companys si són testimonis de la baixa qualitat de la feina feta, com passa quan s'avaluen exposicions orals en un seminari o si els informes de laboratori escrits es presenten públicament en forma de pòster.
- La possible censura dels companys si un dels integrants del grup de treball (que s'avaluarà) no fa correctament la seva part de la feina.

Condició 2

Ocupar els estudiants amb activitats avaluable que els orientin sobre com distribuir correctament el temps i l'esforç en funció dels aspectes més importants de l'assignatura

Aquesta condició té relació amb la direcció i la qualitat de l'esforç, l'orientació que ha de prendre i de quina manera. En general, els estudiants es distribueixen el temps de manera irregular entre les assignatures, sovint concentrant-se en els temes relacionats amb l'avaluació i prou. Si dibuixessin un gràfic amb l'esforç setmanal dedicat durant totes les setmanes del curs a una assignatura que tingués assignada una seqüència d'activitats, el perfil s'assemblaria més als Alps que no pas a Holanda. Els exàmens poden tenir l'efecte de concentrar l'estudi en períodes curts però intensos d'activitats d'estudi al final del curs (com ara revisar apunts) i fins força setmanes després d'acabar l'activitat docent. Activitats de classe habituals (com els fulls d'exercicis o problemes) o proves (com els exàmens informatitzats) poden ajudar els estudiants

a repartir l'esforç al llarg de l'assignatura, tot sovint en una distribució setmanal, mentre que activitats menys freqüents (com els treballs escrits llargs) poden traduir-se en un estudi intensiu d'una setmana o dues just abans de la data de lliurament i, alhora, és possible que temes que no estan presents a les activitats proposades quedin totalment araconats.

De fet, no sabem massa cosa sobre com distribueixen l'esforç els estudiants, i els docents universitaris normalment tenen poca informació sobre què fan els seus estudiants i com ho fan.

Condicció 3

Afrontar una tasca avaluada implica assumir unes activitats d'aprenentatge productives adients

Aquest aspecte afecta els diferents tipus d'estudi i d'activitats d'aprenentatge implicats en la manera d'afrontar una tasca o de preparar-se per a una prova. Hi ha formes d'avaluació que s'associen a activitats d'aprenentatge inútils i inadequades, malgrat que és possible que es tradueixin en qualificacions fiables. Estudiar per a un examen de resposta múltiple pot induir els estudiants a un tipus d'aprenentatge superficial (Scouler i Prosser, 1994; Tang, 1994; Scouler, 1998), igual com els exàmens, malgrat que l'estil d'estudi pugui tenir tanta incidència com la forma de la prova. Els estudiants poden optar per un aprenentatge profund quan es preparen per a proves de resposta múltiple i adoptar estratègies d'estudi encertades, malgrat que la prova plantegi només requisits de nivell baix, i Macdonald (2002) ha indicat que si més no alguns estudiants adoptaven un estil d'aprenentatge profund quan es preparaven per als exàmens i aprenien de manera efectiva, pel fet d'integrar el contingut implicat en la preparació.

L'avaluació falla moltes vegades simplement pel fet que no aconsegueix ocupar els estudiants amb els tipus d'aprenentatge adequats. Lliurar un informe, després de dur a terme al laboratori un procediment dissenyat pel professor, difícilment pot ajudar a l'estudiant a aprendre a dissenyar experiments. Probablement l'única manera d'aprendre a resoldre problemes sigui resolent-ne abans un bon gra-

pat. I probablement l'única manera d'adquirir una certa fluïdesa en el llenguatge de la disciplina sigui practicar una vegada i una altra l'ús d'aquest llenguatge, per exemple a través de l'escriptura. Les activitats i els treballs són el recurs principal per generar aquest tipus de pràctiques. Els estudiants ben poques vegades s'implicaran de debò en pràctiques tan exigents si aquestes activitats no compten de cara a l'avaluació del rendiment o no són, si més no, un requisit de les condicions d'avaluació. No sembla massa probable que els estudiants arribin a redactar treballs llargs i que n'aprenguin les implicacions si no hi estan obligats.

En certa manera, només és feina. Tots aquests treballs, i el pitjor de tot són les lectures, en realitat només són una pila d'hores. (Estudiant d'història) (Hounsell, 1987)

Hi ha formes d'avaluació que poden arribar a orientar malament l'esforç dels estudiants. Snyder (1971) ja va descriure com un grup d'estudiants del Massachussets Institute of Technology, als quals es demanava que fossin imaginatius, deixava de banda qualsevol iniciativa d'aquest tipus quan descobrien que les notes s'obtenien principalment a partir de la pura memorització de matèria per a exàmens de resposta múltiple. Algunes formes d'avaluació es tradueixen en activitats d'aprenentatge adequades, com a efecte colateral. Per exemple, demanar que es redactin treballs tipus assaig pot contribuir al fet que es «llegeixi més enllà» i que s'afavoreixi l'elaboració d'arguments coherents, d'una manera que no seria possible demanant als estudiants simplement que llegissin les lectures recomanades. Si eliminéssim els treballs, mai no s'arribaria a produir la situació d'estudi adequada, ni tan sols en el hipotètic cas que s'aconseguís una quantitat similar de lectures d'un tipus semblant. El producte, el treball, i les notes que hi van associades poden ser menys importants per a l'aprenentatge que el marc que crea l'activitat i que implica activitats d'aprenentatge de l'estil de «llegir més enllà» i «construir arguments». D'una manera semblant, en el cas dels informes de laboratori o d'un *brief* de disseny,² el producte potser és menys important que les particularitats d'estudi que genera aquesta mena de tasca.

2. Document escrit sobre el projecte de disseny que integra el que un client vol d'un dissenyador per a una feina concreta.

Els projectes en grup poden generar més debats entre els estudiants i, per tant, fer que les persones s'enfrontin a punts de vista diferents i que coneguin formes de treballar diferents. La qualitat del producte, fruit del treball en grup (com ara un informe), que es reflecteix a la nota, pot ser menys important que les qualitats del procés d'aprenentatge creats per la tasca.

Tanmateix, els estudiants poden afrontar les tasques dissenyades com a activitats d'aprenentatge de manera que maximitzin les notes obtingudes en comptes de maximitzar l'aprenentatge acumulat amb la implicació en la tasca. D'aquesta manera, és possible que es creï una falsa aparença de coneixement, que una persona es pugui fer passar per competent i coneixedora, i amagar de manera deliberada allò que no entén o desconeix, dient als professors el que volen sentir en comptes de manifestar realment el que pensen, per exemple. En certa manera, aquest fet és el resultat de les preferències de l'estudiant mateix, però les formes d'avaluació, el règim de notes i l'ús del *feedback* poden arribar a ser més determinants que les actituds personals i fomentar fins i tot entre els estudiants un tipus de comportament que en minimitzi l'aprenentatge. En l'exemple següent un estudiant amb una motivació intrínseca descriu en un diari d'aprenentatge la manera com acostuma a fer els treballs d'enginyeria, amb la idea d'obtenir les millors notes en detriment de la qualitat de l'aprenentatge.

El professor normalment vol veure els resultats correctes en un quadre vermell al final del full d'exercicis, si pot ser amb tan poques línies de càlcul com sigui possible: sobretot això, sense cap comentari. Ho odia, es pensa que mires d'omplir la pàgina amb palla perquè sembli que el treball és més llarg. Tampoc no has de deixar-hi els errors, encara que els hagis corregit. Si ho has fet malament, tot llençat. Prefereix pensar que has trobat la solució a la primera. Si continues fent errors, vol dir que no has estudiat prou. No pots tornar a fer un exercici uns mesos més tard perquè només tens els resultats finals, sense comentaris. Si vols provar-ho, potser tornes a fer els mateixos errors que has fet abans, perquè no tens a mà els errors que havies fet abans. (Gibbs, 1992)

La influència del *feedback* en l'aprenentatge

Saber què saps i què no saps ajuda a centrar l'aprenentatge. Els estudiants necessiten un bon *feedback* sobre el seu rendiment per poder treure profit de les assignatures. Quan comencen, els estudiants necessiten ajuda per avaluar quins coneixements tenen i en quin nivell de competència es troben. A les classes, els estudiants han de tenir sovint la possibilitat de rebre suggeriments de millora i poder aplicar-los a la pràctica. En diferents moments dels seus estudis, i quan acaben, els estudiants han de tenir l'oportunitat de reflexionar sobre què han après, què han de continuar aprenent i com s'autoavaluaran. (Chickering i Gamson, 1987)

Tradicionalment, el *feedback* s'ha entès com un aspecte de «correcció d'errors» (Bruner, 1974) o «coneixement de resultats» en relació amb el mateix aprenentatge. Si a un estudiant li diem que va pel bon camí, llavors aprendrà. En aquest article plantegem com pot afectar el *feedback* rebut en el comportament d'aprenentatge de l'estudiant (és a dir, com el *feedback* afecta la presa de decisions de l'estudiant, unes decisions que poden implicar un aprenentatge posterior.

Condicció 4

Un *feedback* suficient, tant pel que fa a la freqüència com al nivell de detall

Aquest aspecte està directament relacionat amb el que normalment definim com a avaluació formativa: l'impacte que pot tenir sobre l'aprenentatge el *feedback* relacionat amb els progressos i que normalment s'ofereix immediatament després d'enllestir una tasca. La quantitat i el nivell de detall del *feedback* poden variar notablement d'una assignatura a una altra, una variació més pronunciada que no pas pel que fa a la quantitat o la qualitat de la docència entre les diferents assignatures.

Aquesta mena de *feedback*, per tal que sigui útil, s'ha d'oferir amb prou regularitat i fer referència a parts relativament petites del contingut de l'assignatura. Un únic *feedback*, per més detallat que sigui, sobre una feina llarga, com ara un treball o un exercici de disseny després

de deu setmanes d'estudi, difícilment pot contribuir a millorar l'aprenentatge durant el curs. Les proves informatitzades han estat un recurs força emprat per tal d'oferir un mínim de *feedback* sobre l'evolució de l'alumne i, en alguns casos, amb l'ajuda de programari desenvolupat específicament per a l'avaluació, hi ha la possibilitat d'oferir un «*feedback* de repàs» si la resposta seleccionada és incorrecta. Cook (2001) va detectar que les notes dels exàmens finals dels estudiants estan estretament relacionades amb el nombre (i, per tant, la freqüència) de les tasques d'avaluació automatitzada que els estudiants havien dut a terme. La freqüència i la rapidesa de resposta d'aquesta mena de *feedback*, que es pot oferir d'una manera raonablement econòmica, pot compensar-ne la qualitat, relativament baixa, i la manca de personalització.

El *feedback*, per tal que pugui ser útil, ha de ser força específic. L'Open University forma els seus 7.500 tutors a temps parcial per tal que puguin oferir un *feedback* detallat i prou ampli. Cole *et al.* (1986) proposen una llista amb les característiques que ha de tenir un bon *feedback* en l'ensenyament a distància, mentre que Roberts (1996) va constatar que les preferències dels estudiants, si més no en relació amb el *feedback*, coincideixen amb aquesta llista. Les formes concretes de *feedback* més efectives varien d'una disciplina a l'altra. A l'obra de Hyland (2001), per exemple, s'hi poden trobar mostres de les formes de *feedback* més efectives en l'aprenentatge d'idiomes. Tant en el camp de la psicologia (Stephenson *et al.*, 1996) com en el de les matemàtiques (Rice *et al.*, 1994), els estudiants han manifestat les seves preferències per un *feedback* específic, detallat i incentivador. Greer (2001) presenta un estudi que explica concretament quina mena d'impacte pot tenir el *feedback* en l'aprenentatge de comptabilitat.

Una bona part del *feedback* que reben els estudiants d'educació superior seria rebutjat pels tutors de personal docent de l'Open University (que fan el seguiment de la qualificació duta a terme pels tutors), atès que el considerarien totalment inadequat i provocaria l'aplicació de mecanismes de garantia de qualitat i de formació del professorat.

Condicció 5

Orientar el *feedback* vers l'actuació dels estudiants, el seu aprenentatge i les accions que poden controlar, en comptes de centrar-lo en els estudiants o les seves característiques

La bibliografia de què disposem relacionada amb l'avaluació formativa diferencia entre el *feedback* que manifesta a l'estudiant que no hi té res a fer, o bé que el situa dins el 10% més baix d'estudiants (en un nivell d'insuficient, per exemple), i el *feedback* que informa a l'estudiant dels punts en què ha fallat i de què pot fer per posar-hi remei. Unes qualificacions sense *feedback* poden ser especialment negatives, però un *feedback* crític centrat en les característiques personals pot desmotivar i incidir negativament en termes del que anomenem «autoeficàcia», és a dir, la sensació de competència dels estudiants. Es tracta d'un aspecte important, ja que l'autoeficàcia manté una relació molt estreta amb la capacitat d'esforç i persistència en les tasques (Schunk, 1984; 1985), té una gran utilitat a l'hora de predir els èxits acadèmics i està vinculada a l'adopció d'un aprenentatge profund (Thomas *et al.*, 1987). En canvi, el *feedback* centrat en el contingut ofereix a l'estudiant opcions per actuar i està menys associat al seu ego, atès que té relació amb la seva actuació i no pas amb la seva persona. Wootton (2002) ha escrit a bastament sobre l'impacte negatiu de l'avaluació sobre els estudiants en situació de «risc» i es planteja si la funció del sistema ha de ser «fomentar l'aprenentatge o bé mesurar el fracàs».

Condicció 6

Oferir el *feedback* a temps per tal que els estudiants el rebin quan encara els importa i quan tenen temps per fer-lo servir de cara a un aprenentatge posterior o per rebre ajuda.

Aquest aspecte ja es va destacar als «set principis de bones pràctiques en educació superior en l'àmbit del grau» (Chickering i Gamson, 1987; 1991) i es basa en un seguit d'estudis sobre la temporització del *feedback* (per obtenir-ne informació resumida, consulteu Dunkin, 1986 i McKeachie *et al.*, 1986). Un mètode docent conegut com a *Personalised System of Instruction* (PSI), que fa un gran èmfasi en el *feedback* immediat en cadascuna de les fases de l'evolució de l'estudiant durant les

unitats de l'assignatura, ha demostrat, segons els indicis de diferents estudis, que pot contribuir a millorar el rendiment de l'estudiant (Kulik *et al.*, 1980).

Si els estudiants no reben el *feedback* de seguida, ja començaran a treballar nous continguts i, quan el rebin, no els servirà per a la feina que estan fent en aquell moment i, per tant, probablement no ajudarà a generar una activitat d'aprenentatge complementària que es tradueixi en resultats. A causa de la pressió provocada per la manca de mitjans, el *feedback* s'ofereix d'una manera més lenta i, atès que al Regne Unit les assignatures ara són més curtes, pot produir-se que un alumne rebi el *feedback* sobre la feina feta quan l'assignatura ja s'hagi acabat. Així, gran part d'aquest *feedback* tan valuós i costós potser serà debades. Per tant, cal trobar un equilibri entre la rapidesa i la qualitat del *feedback*, com passa per exemple quan s'ofereix *feedback* imperfecte a un estudiant que col·labora en tasques de docència, de manera gairebé immediata, una acció que pot traduir-se en un efecte molt més positiu que no pas el *feedback* més complet a càrrec d'un tutor, però que rebrà quatre setmanes després.

Carroll (1995) parla d'uns «tallers d'avaluació formativa» per a grups de 300 estudiants de medicina que constaven de proves de resposta múltiple i, immediatament després, d'una tutoria de «recuperació» sobre el tema. Aquest sistema no oferia un *feedback* individualitzat, però sí immediat, i les sessions del taller es programaven de manera que els estudiants tinguessin l'opció de dedicar temps a estudiar més materials abans de passar a la següent fase del curs. El 85% dels estudiants va manifestar que volia més sessions d'aquesta mena. Sly (1999) documenta l'impacte de les proves com a eina per «practicar» la realització de l'examen posterior. Els estudiants tenien l'opció de fer «pràctiques de test», amb *feedback* informatitzat, i amb prou temps d'antelació abans de l'examen per tal de poder fer servir el *feedback* a l'hora de revisar els punts febles en què havien de millorar. Fins a 197 estudiants fluixos van optar per fer aquestes «pràctiques de test» i tots van millorar la puntuació obtinguda finalment als exàmens, de manera que van arribar a superar a 417 estudiants en principi més bons. Els efectes positius també es van posar de manifest en un examen posterior.

Condicció 7

Un *feedback* adequat en relació amb la finalitat de l'activitat i amb els criteris d'avaluació

Aquest aspecte fa referència a la relació entre el *feedback* i la finalitat d'una activitat d'aprenentatge i els criteris que es fan servir per valorar la realització d'aquesta activitat. El *feedback* pot tenir diferents funcions i, per exemple, es pot emprar fonamentalment per:

- Corregir errors.
- Millorar la comprensió a través d'explicacions.
- Generar més aprenentatge suggerint activitats d'estudi específiques.
- Desenvolupar competències genèriques centrant l'atenció en l'efecte de l'ús d'aquestes activitats i no tant en el contingut.
- Fomentar estratègies metacognitives demanant als estudiants que parin atenció al procés d'aprenentatge vinculat a l'activitat proposada i que l'analitzin.
- Animar els estudiants a continuar estudiant.

A l'hora d'esbrinar quina d'aquestes funcions és la més adequada, cal tenir en compte primer de tot les motivacions d'una activitat concreta. Per exemple, si la intenció era oferir una única oportunitat per practicar l'aplicació d'un procediment o l'ús d'un algoritme d'una manera rigorosa o bé es tractava d'oferir una oportunitat, entre moltes altres, per practicar l'ús d'una competència transversal, o bé es pretenia oferir una oportunitat complexa per reflexionar sobre l'aprenentatge, o bé es buscava plantejar un primer problema senzill en el marc de l'assignatura per motivar l'estudiant a no abandonar i continuar fins al final.

Un estudi recent a l'Open University suggeria que mantenir la motivació era l'aspecte més important i que més podia incidir en els estudiants acabats de matricular a l'hora d'emprendre la primera activitat d'una assignatura (Gibbs i Simpson, 2002). Si un estudiant busca que l'animin i només rep correccions sobre els seus errors, probablement no se'n potenciarà gaire l'aprenentatge.

Els estudiants han d'entendre per què han rebut una determinada qualificació o comentari i per què no n'han rebut un de millor (o més baix). Els criteris han de ser explícits i els estudiants els han d'entendre i els han de copsar nítidament a les seves qualificacions. De vegades els criteris no estan directament vinculats a paràmetres i nivells de rendiment estandaritzats i els estudiants ho tenen difícil per saber quin nivell se n'espera o quin es podria considerar inadequat. Una gran part de la bibliografia sobre l'ús de l'autoavaluació i l'avaluació paritària fa referència a la fiabilitat d'aquesta mena d'avaluacions i admet que l'autoavaluació i l'avaluació paritària constitueixen fonamentalment mecanismes per estalviar feina i personal. Tanmateix, el valor real possiblement s'hagi de buscar en la contribució al fet que els estudiants interioritzin els criteris que se n'espera, de manera que ells mateixos es puguin supervisar i puguin millorar la qualitat de les seves activitats abans de lliurar-les.

Els estudiants han de poder entendre els criteris per tal d'orientar-se correctament en la tasca que se'ls encomani. Penny i Grover (1996) han parlat de com els estudiants poden arribar a malinterpretar els criteris que es faran servir per avaluar el seu projecte de recerca de final de carrera. Els estudiants esperaven criteris relacionats amb objectius de baix nivell, com ara l'estil i la presentació, mentre que els professors feien més èmfasi en objectius d'alt nivell, com la comprensió teòrica i conceptual. Si es generen oportunitats per oferir *feedback* en diferents moments del projecte en marxa és possible reorientar l'esforç dels estudiants d'una manera adequada (Carless, 2002).

L'avaluació, per tant, té un paper molt important a l'hora de consolidar els criteris pels quals s'han de regir els estudiants. Definir les expectatives forma part dels «set principis de bones pràctiques en educació superior en l'àmbit del grau» (Chickering i Gamson, 1987). El *feedback*, les respostes-model i els exemples més destacats (Orsmond *et al.*, 2002) ajuden a reforçar aquestes expectatives i l'autoavaluació ajuda els estudiants a interioritzar-les.

Condicció 8

El *feedback* és adequat en relació amb el que els estudiants creuen que han de fer

Concepcions de la tasca que tenen els estudiants

Els estudiants han d'entendre quina mena de tasca han de dur a terme quan es proposa una activitat d'aprenentatge i què es considerarà «correcte» quan ho intentin. És possible que les seves concepcions siguin equivocades o confuses, independentment de la informació inicial o del *feedback* de què disposin, tal com es pot constatar en aquest exemple:

—Què et sembla que volia buscar el tutor en aquest treball?

Bé... de fet, vaig una mica perdut. Sé que al tutor li agrada que sigui concís, no li agraden les generalitzacions ni els detalls innecessaris, però de tota manera, en general, crec que li agraden més els detalls que les generalitzacions. I com que era un tema tan general, em vaig trobar perdut, perquè no sabia què volia que fes. (Hounsell, 1987)

Independentment del *feedback* que el tutor ofereixi a l'estudiant, sempre s'interpretarà a partir de les idees que l'estudiant s'hagi format sobre què n'esperava el tutor de la seva feina o sobre què cal fer realment. Els estudiants de vegades tenen molts problemes per entendre quin és el sentit d'un treball (com a forma escrita de comunicació en què l'únic lector en sap més que ells sobre el tema) o per a què serveix un informe de laboratori (quan s'ha escrit centenars de vegades en aquell mateix format) o bé quin és l'objectiu d'un projecte de disseny (quan només s'avaluarà el producte final i no pas l'aprenentatge vinculat al procés de creació). A la universitat, hi ha moltes activitats que, per als estudiants, no tenen massa sentit. Aquest fet, inevitablement, provoca problemes a l'hora de llegir el *feedback* i d'entendre si, segons el tutor, han dut a terme correctament aquesta tasca incomprensible.

Concepcions de l'aprenentatge dels estudiants

A més a més de la confusió dels estudiants que hem esmentat al voltant de què n'espera realment el tutor, també podem detectar una concepció *naïve* de l'aprenentatge. Säljö (1982) va observar que els estudiants conceben l'aprenentatge d'una d'aquestes cinc maneres:

1. L'aprenentatge com a recepció passiva d'informació.
2. L'aprenentatge com a memorització activa d'informació.
3. L'aprenentatge com a memorització activa d'informació o procediments per emprar en algun moment del futur.
4. L'aprenentatge com a comprensió.
5. L'aprenentatge com a canvi de la realitat personal: percepció del món de manera diferent.

Un estudiant amb les concepcions dels punts 1, 2 o 3 pot ser que tingui problemes per interpretar un *feedback* que comenti, per exemple, «No hi ha prou discussió crítica» si ha facilitat al tutor una informació compilada de manera diligent i rigorosa. El *feedback* s'ha d'adequar a les concepcions ingènues de l'aprenentatge que de vegades es manifesten en la feina dels estudiants.

Concepcions del coneixement que tenen els estudiants

El «pla de desenvolupament intel·lectual i ètic» de Perry descriu l'evolució dels estudiants durant la seva experiència universitària pel que fa a la comprensió de què és el coneixement (Perry, 1970). Segons aquest autor, els estudiants comencen els estudis pensant que hi ha una gran quantitat de respostes correctes i que la seva feina és aprendre-les i explicar-les correctament al professor. Perry defineix aquest procés d'aprenentatge amb una frase memorable: «addició quantitativa de correcció discreta». Descriu els estudiants com a individus que passen per una sèrie d'estadis de comprensió de la naturalesa del coneixement cada vegada més profunds, entre els quals trobem, per exemple, un relativisme extrem, en què totes les respostes són considerades vàlides. Un estudiant que presenti un treball sense conclusió pot voler transmetre la idea que la deixa a criteri del lector, atès que totes les conclusions tenen la mateixa validesa. En aquest cas, un *feedback* que es limiti a destacar el fet que «falta conclusió» segurament no ajudarà a progressar l'estudiant. El *feedback* dels professors sovint (tot i que no sempre) es genera des d'una posició epistemològica més sofisticada que la de l'estudiant i això fa que sigui molt possible una interpretació incorrecta o, fins i tot, una incomprensió absoluta.

La concepció del llenguatge de la disciplina en l'estudiant

Lea i Street (1998) expliquen el cas d'un estudiant que, després de lliurar un treball d'una assignatura d'història, va rebre el *feedback* següent: «M'agraden les conclusions del teu treball. Són rellevants i estan ben argumentades». Al mateix moment, l'estudiant rebia *feedback* d'un altre treball per a una assignatura paral·lela d'antropologia, on es qüestionava tan durament la seva capacitat de formular arguments clars o de concloure de manera justificada que se li recomanava demanar ajuda als serveis de suport als estudiants. Lea i Street interpreten aquesta situació com una conseqüència de les diferents maneres d'elaborar el discurs en cada disciplina i això, en antropologia, implicava recórrer a diferents formes d'argumentació i a l'ús de proves. Per tant, el problema no era en cap cas l'ús d'una escriptura acadèmica inadequada per a un treball. Si l'estudiant no dominava el discurs de l'antropologia i no l'havia practicat prou, difícilment l'ajuda per escriure textos acadèmics en general podia servir-li d'alguna cosa, independentment de si l'ajuda venia d'un professor o d'un orientador en tècniques d'estudi i d'escriptura acadèmica.

El *feedback* ha de ser diferent en funció del tipus d'escriptura que s'espera i segons el que en sàpiguin els estudiants. En estructures modulars d'assignatures és habitual que els estudiants travessin les fronteres disciplinàries i hagin d'enfrontar-se amb aquest tipus de diferències pel que fa al discurs. Els estudiants de ciències i de carreres tècniques solen tenir problemes amb els treballs típics de les ciències socials, tot i que siguin perfectament capaços d'articular un text a les seves disciplines. Trobem, però, també diferències discursives profundes dins les ciències socials, per exemple, entre la sociologia i la psicologia i dins les humanitats, entre la història i la literatura.

Higgins *et al.* (2001), per la seva banda, investiguen els casos de comunicació fallida del *feedback*. Descriuen un cas en què tot el *feedback* d'un tutor era: «Un esforç suficient. Una anàlisi crítica més profunda de les qüestions clau no hagués anat malament». L'estudiant, que no es conformava amb el «suficient», va quedar frustrat amb aquesta anàlisi crítica tan poc concreta del tutor.

Condicció 9

Recepció i anàlisi del *feedback*

Hi ha diversos estudis que expliquen que els estudiants, quan els tornen el treball corregit, fan una ullada a la nota final i després simplement els llencen a la paperera, juntament amb tot el *feedback*.

A vegades llegeixo de veritat els comentaris, però sé que no tornaré a escriure el mateix text mai més... D'alguna manera tendeixo a ignorar-los, tret que hi posi alguna cosa realment molt alarmant. (Hounsell, 1987)

Crooks (1988) ha sintetitzat un gran nombre d'investigacions sobre aquest tema i ha arribat a la conclusió que en aquells casos en què les qualificacions de proves parcials o treballs de curs tenen un pes considerable en la nota final, els estudiants fan menys cas del *feedback* que reben. Jackson (1995) va observar que els estudiants de tercer any eren especialment propensos a mirar només la nota i fer cas omís dels comentaris dels treballs. També explicava que als estudiants els agrada veure el *feedback*, però sobretot per comprovar que el professor ha llegit el treball amb atenció i l'ha puntuat de manera justa.

No hi ha cap manera de tenir la seguretat que els estudiants llegiran i faran cas del *feedback*, malgrat que el professor l'hagi redactat amb tota cura i que l'hagi tornat amb la màxima rapidesa. Tanmateix, poden seguir una sèrie de passos per aconseguir implicar-nos en el *feedback*:

- Demanar els estudiants que especifiquin sobre quins aspectes volen rebre *feedback* (en relació amb l'activitat proposada) i no donar-los-en sobre res més.
- Donar el *feedback* sense la nota, de manera que tots els estudiants hagin de llegir-se'l per fer-se una idea de com avancen.
- Demanar els alumnes que autoavaluïn les seves activitats (sense condicionar la qualificació), perquè s'adonin de si el seu punt de vista coincideix amb el del professor.
- En una revisió de la bibliografia sobre autoavaluació i avaluació entre iguals, Dochy *et al.* (1999) destaquen que s'ha observat en investigacions controlades que l'avaluació entre iguals millora l'actuació dels

estudiants (en comparació amb un grup de control) i que augmenta el seu domini de les estratègies d'aprenentatge.

- Fer servir activitats en dues fases amb *feedback* al final de la primera, per tal que l'estudiant pugui millorar la qualitat de la feina quan la presenti a la segona fase, que és l'única que compta per a la qualificació. Cooper (2000) ha documentat que aquest sistema pot millorar l'actuació de la majoria d'estudiants, sobretot de la dels més fluixos.
- Donar la qualificació només després de l'autoavaluació i després de rebre el *feedback* del tutor. Taras (2001) descriu l'efectivitat d'una seqüència d'aquesta mena en el marc de l'avaluació sumatòria.

Condicció 10

Resposta de l'estudiant vers el *feedback*

Aquest tema té relació amb l'impacte del *feedback* en l'aprenentatge futur. El *feedback* pot corregir errors de manera precisa, però també és possible que no incideixi en absolut en la manera com un estudiant fa front a l'activitat següent o inicia qualsevol altra tasca d'aprenentatge. Aquesta situació es pot produir per diversos motius:

- El *feedback* de vegades arriba massa tard i els estudiants no poden posar-lo en pràctica.
- El *feedback* sovint mira només cap enrere, fent referència a qüestions vinculades a una matèria que no tornarà a ser objecte d'estudi, en lloc de mirar cap endavant i encarar les activitats d'estudi que vindran o les tasques que l'estudiant haurà de dur a terme.
- El *feedback* és poc realista o poc precís pel que fa als esforços que demana a l'estudiant (per exemple, «Llegeix la bibliografia» en lloc de «Consulta Smith cap. 2, p. 24-29, on s'explica una concepció totalment diferent»).
- El *feedback* demana a l'estudiant que faci coses que no sap com fer (per exemple, «sigues més sociològic» o «expressa't amb més claredat»).
- El *feedback* està molt vinculat al context i només pot aplicar-se a una activitat concreta en comptes d'abordar aspectes més generals, com ara maneres de treballar o perspectives rellevants per a d'altres activitats.

- El *feedback* pot desanimar i fer que l'estudiant s'esforci menys en els estudis en lloc d'abocar-s'hi més.
- No es fa un seguiment per comprovar que els estudiants han pres mesures a partir del *feedback*, de manera que se'ls permet que ignoren el *feedback* amb total impunitat.

Ding (1998) afirma que, encara que els estudiants llegeixin els comentaris del *feedback*, en fan molt poca cosa. En canvi, Brookheart (2001) apunta que els estudiants amb bons resultats ho fan servir tot activament (qualificacions, *feedback* i autoavaluació) per aprendre i orientar les futures activitats d'estudi. Les variables més importants són, com passa sovint, les que tenen a veure amb l'estudiant més que amb el professor. Ensenyar els estudiants a supervisar-se la feina és, segons l'anàlisi teòrica de Sadler (1989) sobre el paper del *feedback*, l'objectiu final del *feedback*.

Diverses investigacions sobre l'ús de l'anomenada «avaluació a la classe» en universitats nord-americanes evidencien l'impacte d'aquesta avaluació, no en l'aprenentatge d'un contingut específic sinó en el desenvolupament metacognitiu dels estudiants i la seva capacitat d'incrementar el control sobre l'aprenentatge (vegeu Steadman, 1998 per consultar-ne un resum). Semblaria, doncs, que cal ensenyar els estudiants a fer servir el *feedback* per desenvolupar el control metacognitiu (Sadler, 1989). Ara bé, la millora de la capacitat d'aprenentatge pot tenir uns efectes inesperats. Ramsden *et al.* (1987), analitzant un programa de «tècniques d'estudi» dissenyat per aconseguir incrementar el nombre d'estudiants que adopten un aprenentatge en profunditat, van observar com en realitat passava tot el contrari. L'augment de la consciència sobre la forma d'aprendre permetia als estudiants adoptar una forma d'aprenentatge més superficial, ja que s'adaptaven al baix nivell que percebien als procediments d'avaluació de les assignatures que cursaven! Això ens torna a demostrar que la manera com perceben l'avaluació els estudiants influeix en l'aprenentatge.

CONCLUSIÓ

Les condicions d'avaluació que afavoreixen l'aprenentatge són objecte d'un estudi pràctic en el marc d'un projecte a gran escala que s'ha iniciat amb un seguiment de l'avaluació aplicada a assignatures de ciències de dues universitats. Els professors de les diferents assignatures, en què s'utilitza un ampli ventall de formes d'avaluació, recullen dades sobre, per exemple, com distribueixen els estudiants l'esforç en relació amb el que se'ls demana a l'avaluació i com responen al *feedback*. Aleshores fan servir aquesta informació per diagnosticar problemes potencials a les seves assignatures, introduir canvis a l'avaluació per resoldre'ls i, finalment, avaluar si els canvis han repercutit positivament en la manera d'aprendre dels estudiants. Aquest sistema és molt similar a qualsevol procés d'investigació-acció per tal de millorar la docència, però amb una diferència cabdal: l'objectiu no és la metodologia d'ensenyament sinó l'avaluació. La hipòtesi de partida és que hi ha més marge de millora en la docència canviant aspectes de l'avaluació que canviant qualsevol altra cosa i, alhora, que els professors saben menys sobre com responen els estudiants a l'avaluació que sobre qualsevol altre aspecte. A mesura que avanci el projecte, la comprensió dels professors i les proves sobre els canvis més efectius introduïts a les assignatures permetran desenvolupar més aquestes «condicions». Es tracta d'una tasca col·laborativa a gran escala sobre el «coneixement de l'avaluació» que proporcionarà estudis de cas sobre els canvis que han estat efectius i també un marc conceptual elaborat que ajudarà a explicar per què ho han estat. L'objectiu és que aquestes condicions formin part d'una llista de control a l'abast de qualsevol professor que vulgui revisar l'efectivitat del sistema d'avaluació de les seves assignatures pel que fa a l'aprenentatge.

AGRAÏMENTS

El buidatge bibliogràfic que configura la base d'aquest article hem d'agrair-lo a l'Student Support Research Group de l'Open University. El projecte de recerca que ha motivat el buidatge bibliogràfic, «Millora de l'efectivitat de l'avaluació formativa en ciències», ha estat possible gràcies a la dotació de 250.000 lliures atorgada pel Higher Education Funding Council for England dins del Programa per al Desenvolupament de la Docència i l'Aprenentatge.

BIBLIOGRAFIA

- ANGELO, T.A. i CROSS, K.P. (1993). *Classroom Assessment Techniques: a handbook for college teachers*. San Francisco (CA): Jossey-Bass.
- BAIRD, L.L. (1985). «Do grades and tests predict adult accomplishment?». *A Research in Higher Education*, 23 (1), 3-85.
- BERLINER, D.C. (1984). «The half-full glass: a review of research on teaching». P.L. Hosford (ed.). *Using What We Know About Teaching*. Alexandria (VA): Association for Supervision and Curriculum Development.
- BIGGS, J.B. i COLLIS K.F. (1982). *Evaluating the Quality of Learning: the SOLO taxonomy*. Nova York: Academic Press.
- BLACK, P. i WILIAM, D. (1998). «Assessment and classroom learning». *A Assessment in Education*, 5 (1), 7-74.
- BRIDGES, P., COOPER, A., EVANSON, P., HAINES, C., JENKINS, D., SCURRY, D., WOOLF, H. i YORKE, M. (2002). «Coursework marks high, examination marks low: discuss». *A Assessment and Evaluation in Higher Education*, 27 (1), 36-48.
- BRITTINGHAM, B.E. (1988). «Undergraduate students' use of time: a classroom investigation». *A To Improve the Academy*, 7, 45-52.
- BROOKHEART, S. M. (2001). «Successful students' formative and summative uses of assessment information». *A Assessment and Evaluation in Higher Education*, 8 (2), 154-169.
- BRUNER, J.S. (1974). *Toward a Theory of Instruction*. Cambridge (MA): Harvard University Press.
- CARLESS, D.M. (2002). «The 'mini-viva' as a tool to enhance assessment for learning». *A Assessment and Evaluation in Higher Education*, 27 (4), 353-363.
- CARROLL, M. (1995). «Formative assessment workshops: feedback sessions for large classes». *A Biomedical Education*, 23 (2), 65-67.
- CHANSARKAR, B.A. i RAUT-ROY, U. (1987). «Student performance under different assessment situations». *A Assessment and Evaluation in Higher Education*, 12 (2), 115-122.
- CHICKERING, A.W. i GAMSON, Z.F. (1987). *Seven Principles to Good Practice in Undergraduate Education*. Racine (WI): The Johnson Foundation Inc.
- (1991). *Applying the Seven Principles to Good Practice in Undergraduate Education*. San Francisco: Jossey-Bass.

- COLE, S., COATES, M. i LENTELL, H. (1986). «Towards good teaching by correspondence». *A Open Learning*, 1 (1), 16-22.
- CONWAY, M.A., COHEN, G. i STANHOPE, N. (1992). «Why is it that university grades do not predict very long-term retention?» *A Journal of Experimental Psychology: General*, 121 (3), 49-57.
- COOK, A. (2001). «Assessing the use of flexible assessment». *A Assessment and Evaluation in Higher Education*, 26 (6), 539-549.
- COOPER, N.J. (2000). «Facilitating learning from formative feedback in level 3 assessment». *A Assessment and Evaluation in Higher Education*, 25 (3), 279-291.
- CROOKS, T.J. (1988). «The impact of classroom evaluation practices on students». *A Review of Educational Research*, 58 (4), 438-481.
- CURTIS, S. i SHAMI, N. (2002). «The effect of taking paid employment during termtime on students' academic studies». *A Journal of Further and Higher Education*, 26 (2), 129-138.
- DING, L. (1998). *Revisiting assessment and learning: implications of students' perspectives on assessment feedback*. Ponència presentada al congrés anual de la Scottish Educational Research Association, Universitat de Dundee, 25-26 setembre.
- DOCHY, F., SEGERS, M. i SLUIJSMANS, D. (1999). «The use of self-, peer- and coassessment: a review». *A Studies in Higher Education*, 24 (3), 331-350.
- DUNKIN, M.J. (1986). «Research on Teaching in Higher Education». A M.C. Wittrock (ed.). *Handbook of Research on Teaching*, Nova York: Macmillan 3a ed.
- FORBES, D. i SPENCE, J. (1991). «An experiment in assessment for a large class». A R. Smith (ed.). *Innovations in Engineering Education*. Londres: Ellis Horwood.
- GAGNE, R.M. (1977). *The Conditions of Learning*. Nova York: Holt, Rinehart and Winston (3a ed.).
- GARDINER, L.F. (1997). «Redesigning higher education: producing dramatic gains in student learning». *A ASHE-ERIC Higher Education Reports*, 23 (7), Washington (DC): Association for the Study of Higher Education.
- GIBBS, G. (1992). *Assessing More Students*. Oxford: Oxford Centre for Staff Development.
- (1995). *Assessing Student Centred Courses*. Oxford: Oxford Centre for Staff Development.
- (1999). «Using assessment strategically to change the way students learn». A S. Brown i A. Glasner (ed.). *Assessment Matters in Higher Edu-*

- cation. Buckingham: Society for Research into Higher Education and Open University Press.
- GIBBS, G. i LUCAS, L. (1987). «Coursework assessment, class size and student performance: 1984-94». A *Journal of Further and Higher Education*, 21 (2), 183-192.
- GIBBS, G. i SIMPSON, C. (2002). *Evaluation of Regional Retention Activity: interim report*. Open University: Student Support Research Group Report 40/2002.
- GREER, L. (2001). «Does changing the method of assessment of a module improve the performance of a student?» A *Assessment and Evaluation in Higher Education*, 26 (2), 128-138.
- HATTIE, J.A. (1987). «Identifying the salient facets of a model of student learning: a synthesis of meta-analyses». A *International Journal of Educational Research*, 11, 187-212.
- HIGGINS, R., HARTLEY, P. i SKELTON, A. (2001). «Getting the message across: the problem of communicating assessment feedback». A *Teaching in Higher Education*, 6 (2), 269-274.
- HOUNSELL, D. (1987). «Essay writing and the quality of feedback». A J.T.E. Richardson, M.W. Eysenck i D. Warren-Piper (ed.). *Student Learning: research in education and cognitive psychology*. Milton Keynes: Open University Press and Society for Research into Higher Education.
- HUTCHINGS, P., MARCHESE, T. i WRIGHT, B. (1991). *Using Assessment to Strengthen General Education*. Washington (DC): American Association for Higher Education.
- HYLAND, F. (2001). «Providing effective support: investigating feedback to distance language learners». A *Open Learning*, 16 (3), 231-247.
- INNIS, K. (1996). *Diary Survey: how undergraduate full-time students spend their time*. Leeds: Leeds Metropolitan University.
- JACKSON, M. (1995). *Making the grade: the formative evaluation of essays* [en línia]. UtiliBASE: disponible a <http://ultibase.rmit.edu.au/Articles/dec96/jacks1.htm>.
- JAMES, D. i FLEMING, S. (2004-05). «Agreement in Student Performance in Assessment». A *Learning and Teaching in Higher Education*, 1, 32-50.
- KEMBER, D., NG, S., TSE, H., WONG, E.T.T. i POMFRET, M. (1996). «An examination of the interrelationships between workload, study time, learning approaches and academic outcomes». A *Studies in Higher Education*, 21 (3), 347-358.

- KNIVETON, B.H. (1996). «Student Perceptions of Assessment Methods». A *Assessment & Evaluation in Higher Education*, 21 (3), 229-238.
- KULIK, C.C., KULIK, J.A. i COHEN, P.A. (1980). «Instructional technology and college teaching». A *Teaching Psychology*, 7, 199-205.
- LEA, M. i STREET, B. (1998). «Student writing in higher education: an academic literacies approach». A *Studies in Higher Education*, 23(2), 157-172.
- MACDONALD, J. (2002). «Getting it together and being put on the spot»: synopsis, motivation and examination». A *Studies in Higher Education*, 27 (3), 329-338.
- MACLELLEN, E. (2001). «Assessment for learning: the different perceptions of tutors and students». A *Assessment and Evaluation in Higher Education*, 26 (4), 307-318.
- MACFARLANE, B. (1992). «The “Thatcherite” generation of university degree results». A *Journal of Further and Higher Education*, 16, 60-70.
- MARTON, F. i WENESTAM, C.G. (1978). «Qualitative differences in the understanding and retention of the main points in some texts based on the principle-example structure». A M.M. Gruneberg, P.E. Morris i R.N. Sykes (ed.). *Practical Aspects of Memory*, Londres: Academic Press.
- MCKEACHIE, W.J., PINTRICH, P.R., LIN, Y. i SMITH, D. (1986). *Teaching and Learning in the College Classroom: a review of the research literature*. Ann Arbor: National Centre for Research to Improve Postsecondary Teaching and Learning, Universitat de Michigan.
- MELTON, R. (2000). *Assessment and Assignments — a review of related issues based on data collected from the 1998 courses survey*, Report SRC 200, Open University.
- MENTKOWSKI, M., ASTIN, A., EWELL, P. i MORAN, E. T. (1991). *Catching Theory up with Practice: conceptual frameworks for assessment*. Washington (DC): AAHE Assessment Forum.
- MILLER, C.M.I. i PARLETT, M. (1974). *Up to the Mark: a study of the examination game*. Guildford: Society for Research into Higher Education.
- MOFFAT, M. (1989). *Coming of Age in New Jersey: college and American culture*. New Brunswick (NJ): Rutgers University Press.
- ORSMOND, P., MERRY, S. i REILING, K. (2002). «The use of exemplars and formative feedback when using student derived marking criteria in peer and self assessment». A *Assessment and Evaluation in Higher Education*, 27 (4), 309-323.

- PATON-SALTZBERG, R. i LINDSAY, R. (1993). *The Effects of Paid Employment on the Academic Performance of Full-time Students in Higher Education*. Oxford: Oxford Polytechnic.
- PENNY, A.J. i GROVER, C. (1996). «An analysis of student grade expectations and marker consistency». *A Assessment and Evaluation in Higher Education*, 21 (2), 173-184.
- PERRY, W.G. (1970). *Forms of Intellectual and Ethical Development in the College Years*. Nova York: Holy, Rinehart and Winston.
- RAMSDEN, P. (1991). «A performance indicator of teaching quality in higher education: the Course Experience Questionnaire». *A Studies in Higher Education*, 16, 129-150.
- (1992). *Learning to teach in Higher Education*. Londres: Routledge.
- RAMSDEN, P., BESWICK, D. i BOWDEN, J. (1987). «Learning processes and learning skills». A J.T.E. Richardson, M.W. Eysenck i D.W. Warren-Piper (eds.). *Student Learning: research in education and cognitive psychology*. Milton Keynes: Open University Press and Society for Research into Higher Education.
- RICE, M., MOUSLEY, J. i DAVIS, R. (1994). «Improving student feedback in distance education: a research report» A T. Evans i D. Murphy (eds.). *Research in Distance Education 3: revised papers from the third research in distance education conference*. Geelong, Victoria: Deakin University Press.
- ROBERTS, D. (1996). «Feedback on assignments». *A Distance Education*, 17 (1), 95-116.
- ROWNTREE, D. (1987). *Assessing Students — how shall we know them?* Londres: Kogan Page.
- SADLER, D.R. (1989). «Formative assessment and the design of instructional systems». *A Instructional Science*, 18, 119-144.
- (1998). «Formative assessment: revisiting the territory». *A Assessment in Education*, 5 (1), 77-84.
- SÄLJÖ, R. (1982). *Learning and Understanding*. Göteborg: Acta Universitatis Gothoburgensis.
- SAMBELL, K. i MCDOWELL, L. (1998). «The construction of the hidden curriculum: messages and meanings in the assessment of student learning». *A Assessment and Evaluation in Higher Education*, 23 (4), 391-402.
- SCHUNK, D. (1984). «Self-efficacy perspective on achievement behaviour». *A Educational Psychologist*, 19, 48-58.
- (1985). «Self-efficacy and classroom learning». *A Psychology in the Schools*, 22, 208-223.

- SCOULER, K. (1998). «The influence of assessment method on students' learning approaches: multiple choice question examinations vs. essay assignment». A *Higher Education*, 35, 453-472.
- SCOULER, K. i PROSSER, M. (1994). «Students' experiences of studying for multiple choice question examinations». A *Studies in Higher Education*, 19 (3), 267-280.
- SNYDER, B.R. (1971). *The Hidden Curriculum*. Cambridge (MA): MIT Press.
- SLY, L. (1999). «Practice tests as formative assessment improve student performance on computer managed learning assessments». A *Assessment and Evaluation in Higher Education*, 24 (3), 339-344.
- STARR, J.W. (1970). «Student opinion on methods of assessment». A *Educational Review*, 22, 243-253.
- STEADMAN, M. (1998). «Using classroom assessment to change both teaching and learning». A *New Directions for Teaching and Learning*, 75, 23-35.
- STEPHENSON, K., SANDER, P. i NAYLOR, P. (1996). «Student perceptions of the tutor's role in distance learning». A *Open Learning*, 11 (1), 22-30.
- TANG, C. (1994). «Effects of modes of assessment on students' preparation strategies». A G. Gibbs (ed.). *Improving Student Learning: theory and practice*. Oxford: Oxford Centre for Staff Development, 151-170.
- TARAS, M. (2001). «The use of tutor feedback and student self-assessment in summative assessment: towards transparency for students and for tutors». A *Assessment and Evaluation in Higher Education*, 26 (6), 605-614.
- THOMAS, J.W., IVENTOSH, L. i ROHWER, W.D. (1987). «Relationships among student characteristics, study activities and achievement as a function of course characteristics». A *Contemporary Educational Psychology*, 12, 344-364.
- TINTO, V. (1993). *Leaving College: rethinking the causes and cures of student attrition*. Chicago: University of Chicago Press (2a ed.).
- TORRANCE, H. i PRYOR, J. (1998). *Investigative Formative Assessment: teaching, learning and assessment in the classroom*. Buckingham: Open University Press.
- TYNJALA, P. (1998). «Traditional studying for examination vs constructivist learning tasks: do learning outcomes differ?» A *Studies in Higher Education*, 23 (2), 173-191.
- VOS, P. (1991). *Curriculum Control of Learning Processes in Higher Education*. Edimburg: 13th International Forum on Higher Education of the European Association for Institutional Research.

- WARREN, J.R. (1971). *College Grading Practices: an overview*, Report núm. 9. Washington (DC): ERIC Clearinghouse on Higher Education.
- WOOTTON, S. (2002). «Encouraging learning or measuring failure?» *A Teaching in Higher Education*, 7 (3), 353-357.
- WOTJAS, O. (1998). «Feedback? No, just give us the answers». *A Times Higher Education Supplement*, 25 settembre de 1998.
- YORKE, M. (2001). «Formative assessment and its relevance to retention. *A Higher Education Research and Development*, 20 (2), 115-126.

**CONDICIONS PER A UNA
AVALUACIÓ CONTINUADA QUE
AFAVOREIXI L'APRENENTATGE**

Graham Gibbs fou director de l'Oxford Learning Institute de la Universitat d'Oxford fins al 2007, institut al qual continua vinculat com a investigador visitant. Ha dirigit també l'Student Support Research Group de l'Open University. Abans havia codirigit el Centre per a la Pràctica a l'Ensenyament Superior de l'Open University i havia estat professor i director acadèmic del Centre de Desenvolupament i Aprenentatge d'Oxford Brookes. Des l'any 1997 treballa en el desenvolupament i la implantació d'estratègies de docència i aprenentatge a l'Equip de Coordinació Nacional per al Desenvolupament de l'Ensenyament i l'Aprenentatge de la HEFCE (Higher Education Funding Council for England). És autor de nombroses publicacions sobre docència, aprenentatge i avaluació a l'educació superior i membre cofundador de l'Improving Student Learning Symposia i de l'International Consortium for Educational Development in Higher Education (ICED). A més a més, ha treballat com a assessor en més de 120 institucions d'educació superior d'arreu del món. Ha rebut diverses distincions com a doctor *honoris causa*, la més recent a la Universitat d'Utrecht, per la seva contribució al desenvolupament de la docència i l'aprenentatge a l'educació superior.

Claire Simpson té un màster en mètodes de recerca psicològica i és investigadora de l'Student Support Research Group de l'Open University, responsable dels serveis d'avaluació de la Business School de l'Open University (2003) i assessora en mètodes de recerca, disseny de qüestionaris, selecció de mostres, anàlisi i processament de dades per a projectes relacionats amb el suport a l'estudiant.

Ha col·laborat en l'avaluació del projecte de retenció d'estudiants, la recerca sobre els diferents tipus de *feedback* que rep el docent i l'impacte en l'aprenentatge de l'estudiant. També fa tasques de formació en el camp de la recerca educativa.