



Què funciona en educació?

Evidències per a la millora educativa

9

maig de 2018

És l'avaluació de l'alumnat un mecanisme de millora del rendiment escolar?

Sheila González Motos

Tradicionalment, l'avaluació s'ha entès com la valoració del professorat sobre els aprenentatges de l'alumnat. Avui en dia, però, l'avaluació ha deixat de ser un instrument finalista (atorgament d'una nota) i s'ha convertit en quelcom més complex, en un recurs d'aprenentatge més. Ara bé, l'amplitud i la varietat d'instruments avaluadors aconsella investigar i reflexionar sobre quines pràctiques tenen un impacte més gran. Aquesta revisió de l'evidència pretén aportar dades que permetin saber si l'avaluació formativa és un instrument de millora del rendiment dels estudiants, quins mecanismes avaluadors són més efectius en la millora dels resultats acadèmics i quins perfils d'alumnes es beneficien més de la implementació de les diferents estratègies d'avaluació a l'aula.

“Durant massa temps l'educació s'ha basat en inèrcies i tradicions, i els canvis educatius, en intuïcions o creences no fonamentades. El moviment ‘Què funciona’ irromp en el món de l'educació amb un objectiu clar: promoure polítiques i pràctiques educatives basades en l'evidència. [Ivàlua](#) i la [Fundació Jaume Bofill](#) s'alien per fer avançar aquest moviment a casa nostra.”



Què funciona en educació?

Evidències per a la millora educativa

És l'avaluació de l'alumnat un mecanisme de millora del rendiment escolar?



Sheila González Motos

Doctora en Ciències Polítiques i Transformació Social. Ha fet recerca en l'àmbit de l'educació i la immigració, amb un interès especial en temes de desigualtats socials. Actualment és investigadora postdoctoral i professora al Departament de Sociologia de la Universitat Autònoma de Barcelona.

Motivació

L'avaluació dels aprenentatges forma part de la tasca de tot docent i esdevé una peça central en el sistema educatiu, ateses les implicacions que deriven de l'obtenció de determinades qualificacions, titulacions o acreditacions per als alumnes. Al mateix temps, i amb independència dels instruments emprats, es coincideix a ubicar l'avaluació entre les tasques a què dediquen més temps els docents, circumstància que aconsella identificar quines formes avaluadores són més eficients i suposen una manera millor d'emprar el temps dedicat a l'avaluació. A més a més, en un context com l'actual, en què guanyen força els plantejaments pedagògics que qüestionen l'avaluació com a peça central del sistema escolar, és convenient reflexionar sobre el valor formatiu de diferents mecanismes avaluadors.

Tradicionalment, l'avaluació s'ha entès com la valoració que el professorat fa sobre els aprenentatges de l'alumnat al final d'un període més o menys ampli (unitat didàctica, semestre o curs), sigui a través de proves escrites, exàmens orals, treballs individuals o activitats grupals, entre altres instruments. Els diferents corrents pedagògics, no obstant, hi han anat incorporant elements nous, i d'aquesta manera han convertit l'avaluació en quelcom més complex: no sempre es produeix al final d'un període, sinó que pot ser també inicial (diagnosi) o de progrés (contínua); no sempre

la fa el professorat, sinó que també pot recaure en els companys (coavaluació) o en organismes externs a l'escola (proves de competències bàsiques, PISA, etc.) i, fins i tot, en l'alumne mateix (autoavaluació), i, el més important, l'avaluació deixa de ser un instrument finalista (atorgament d'una nota) i es converteix en un recurs d'aprenentatge més.

És en aquest sentit que l'OCDE va dur a terme una recerca extensa sobre l'impacte dels mètodes d'avaluació desenvolupats a la major part de països, ja no només amb voluntat finalista, sinó com a instruments d'aprenentatge [1]. S'ha escrit molt sobre aquest tema, però hi ha pocs estudis rigorosos que permetin assenyalar una clara línia d'intervenció en matèria avaluado-

ra. La qualitat de l'evidència és, en general, baixa, i les conclusions a què arriba no són unívocues [1] [2]. Malgrat això, algunes recerques aporten dades i reflexions interessants per al tema que ens ocupa. Aquesta revisió de l'evidència posa el focus d'atenció en aquells instruments avaluadors que tenen vocació formativa, és a dir, que no es limiten a l'atorgament d'una nota, sinó que esdevenen part del procés d'aprenentatge. Es tracta, doncs, d'abordar l'impacte sobre el rendiment escolar i no de trobar els mecanismes més justos per certificar els coneixements apresos.

Aquesta revisió de l'evidència posa el focus d'atenció en aquells instruments avaluadors que tenen vocació formativa, és a dir, que no es limiten a l'atorgament d'una nota, sinó que esdevenen part del procés d'aprenentatge. Es tracta, doncs, d'abordar l'impacte sobre el rendiment escolar i no de trobar els mecanismes més justos per certificar els coneixements apresos.



De quins mètodes avaluadors parlem?

L'avaluació, com dèiem, és un àmbit molt ampli que inclou multitud d'instruments i de finalitats. D'entrada, i de forma molt genèrica, podem identificar tres tipus d'avaluacions. En primer lloc, l'avaluació formativa (*formative assessment*), emprada pel professorat i pels alumnes durant el procés d'aprenentatge per tal d'anar ajustant les dinàmiques d'ensenyament i d'estudi. En segon lloc, l'avaluació sumatòria o final (*summative assessment*), que té lloc un cop es tanca una unitat, un semestre o un curs i té per objectiu certificar l'assoliment d'un cert nivell d'aprenentatge (en funció de diferents criteris). I, finalment, l'avaluació estandarditzada (*interim assessment*), que té per objectiu aportar informació comparable entre classes, escoles, districtes o països sobre el grau de rendiment o coneixement de l'alumnat, al servei dels educadors i educadores i dels responsables polítics.

No es tracta, no obstant, de lògiques avaluadores excloents o incompatibles, atès que tant l'avaluació final com l'estandarditzada poden ser emprades amb finalitats formatives o a la inversa, mecanismes dissenyats per a l'avaluació formativa es desprenen del seu valor formatiu quan els i les docents es limiten a atorgar una puntuació o nota [3][4].

Taula 1.
Tipus d'avaluacions

Tipus d'avaluacions	Objectiu	Destinatari	Moment
Avaluació formativa	(Re)orientar el procés d'ensenyament-aprenentatge	Alumnat i docents	Durant el procés d'aprenentatge
Avaluació final o sumatòria	Identificar assoliment d'aprenentatge	Alumnat	Al final del procés (unitat, semestre, curs, etc.)
Avaluació estandarditzada	Comparar assoliments entre classes, escoles, districtes, països...	Docents i responsables polítics	Indiferent

Font: Elaboració pròpia

Com avançàvem, aquesta revisió se centra en l'avaluació formativa o *feedback* formatiu.¹ No hi ha una definició única i completa del que s'entén per avaluació formativa. La inexactitud en l'ús del terme ha generat una ingent literatura, però també una manca d'acord respecte dels

seus resultats. Podem afirmar que la vaguetat en les definicions ha contribuït a la debilitat en els resultats empírics i ha dificultat la identificació de bones pràctiques [3]. Malgrat això, hi ha un consens per identificar com a avaluació formativa aquells instruments emprats en l'àmbit educatiu que responen a un doble objectiu: d'una banda, millorar el coneixement i l'adquisició d'habilitats i, de l'altra, incrementar la motivació per l'aprenentatge [5]. En aquest marc s'encabeixen principalment els instruments avaluadors dissenyats amb aquesta doble finalitat, però també hi ha espai per a aquells altres mètodes que, partint d'altres objectius més propis de l'avaluació estandarditzada o de l'avaluació sumatòria, desenvolupen el seu potencial formatiu.

Hi ha un consens per identificar com a avaluació formativa aquells instruments emprats en l'àmbit educatiu que responen a un doble objectiu: d'una banda, millorar el coneixement i l'adquisició d'habilitats i, de l'altra, incrementar la motivació per l'aprenentatge.



Aquests dos objectius (coneixement i motivació) es poden assolir a partir de diferents models d'avaluació, a priori contraposats, però complementaris a la pràctica:

- L'avaluació pot ser **normativa** (*normative feedback*), si es compara el rendiment de l'alumnat respecte de la resta de companys del grup classe, o **autoreferencial** (*self-reference feedback*), quan el seu rendiment es mesura respecte de les seves capacitats.
- El missatge que acompanya el *feedback* permet distingir entre el **retorn directiu**, és a dir, indicacions a l'alumne sobre què cal que revisi, i el **retorn facilitador**, un conjunt de comentaris i suggeriments que l'ajuden a orientar el seu propi procés de millora.
- En uns casos, l'avaluació se centra en l'exercici presentat per un alumne, de manera que la informació de correcció que facilita el professorat s'adreça a cada alumne sobre l'exercici concret. En altres casos, en canvi, el retorn no es produeix de forma

¹ Malgrat que hi ha autors que identifiquen diferències entre ambdós conceptes, bona part de la literatura revisada els fa servir com a sinònims, indistintament.

individual sinó agregada, és a dir, es presenten davant tota la classe les valoracions sobre el desenvolupament general del grup en una prova feta individualment.

- Una quarta categoria permet diferenciar una **avaluació orientada a l'aprenentatge** (*learning orientation*), que informa l'alumne sobre com de prop o lluny es troba dels objectius finals marcats per al seu aprenentatge, i una **avaluació orientada al rendiment** (*performance orientation*), que es limita a informar-lo dels resultats de les proves que ha fet.
- Si bé la major part de l'avaluació en el sistema escolar la fa el personal docent, hi ha altres fórmules avaluadores, com l'**autoavaluació** (*self-assessment*) o l'**avaluació entre estudiants o coavaluació** (*peer assessment*). La primera avaluació es fonamenta en la capacitat de l'alumnat per entendre els objectius de l'ensenyament, avaluar les tasques fetes, detectar els errors i corregir-los. La coavaluació comporta la revisió d'un treball per part dels companys, a partir d'una sèrie de criteris avaluadors prèviament discutits en grup per tal d'establir una perspectiva comuna d'avaluació.

Taula 2.
Tipus d'avaluacions formatives

Tipus d'avaluació
Avaluació normativa / Avaluació autoreferencial
Avaluació directiva / Avaluació facilitadora
Avaluació individual / Avaluació de grup
Avaluació orientada a l'aprenentatge / Avaluació orientada al rendiment
Avaluació estàndard / Autoavaluació / Avaluació entre estudiants

Font: Elaboració pròpia

Al marge d'aquests models, s'ha de tenir en consideració que el ventall d'objectes per avaluar i sobre els quals s'ha de fonamentar el retorn formatiu és molt ampli. Hattie i Timperley [6] identifiquen tres nivells sobre els quals es pot aplicar el *feedback* formatiu:

- **Tasca concreta desenvolupada per l'alumnat:** es tracta d'indicar el grau d'acompliment de l'estudiant en un exercici concret, tot indicant les respostes correctes i errònies.
- **Procés d'aprenentatge:** en aquest nivell, l'avaluació sobrepassa la valoració de l'exercici concret i s'adreça al procés d'aprenentatge que connecta les diferents tasques que cal fer, tot aportant estratègies per a la detecció d'errors i cerca de solucions.
- **Capacitat d'autoregulació:** implica incloure en l'avaluació aspectes vinculats al compromís i la capacitat d'organització de l'alumne a partir d'indicacions sobre com millorar l'autonomia, la disciplina i l'autocontrol amb vista a assolir l'objectiu d'aprenentatge acordat.

La literatura revisada es fa ressò d'aquest ampli ventall d'opcions pel que fa a l'avaluació. A més a més, una part important incorpora a l'anàlisi altres variables amb què l'avaluació interacciona, com pot ser el nivell acadèmic dels i les estudiants, l'edat o el tipus d'estudis.

Preguntes que guien la revisió

L'avaluació ha estat tradicionalment un dels eixos organitzatius del sistema escolar i, per tant, de la tasca docent. En els darrers anys es produeixen dos processos paral·lels i antagònics: d'una banda, institucions internacionals i poders públics reclamen disposar de més informació sobre el rendiment als centres escolars a través de la introducció de proves diagnòstiques noves i, de l'altra, han agafat força els moviments pedagògics que qüestionen la utilitat de les avaluacions (tant externes com internes). Aquesta revisió de l'evidència pretén aportar dades a aquest debat, tot donant resposta a les següents preguntes:

1. Pot ser l'avaluació formativa un instrument de millora del rendiment escolar dels estudiants?
2. Quins mecanismes avaluadors es mostren més efectius en la millora dels resultats acadèmics? I d'altres competències?
3. Quins perfils es beneficien més de la implementació de diferents estratègies d'avaluació a l'aula?
4. En clau d'intervenció pública, és recomanable apostar per la inversió en avaluació? Sota quins paràmetres?

Revisió de l'evidència

Revisions i estudis considerats

Malgrat la ingent producció acadèmica que des de diferents àmbits (educació, psicologia, sociologia, etc.) s'ha generat entorn de l'avaluació en l'àmbit escolar, hi ha poca evidència robusta que permeti extreure'n conclusions clares. Aquest estudi es fonamenta en les revisions, metanàlisis i estudis experimentals recollits a la [taula 3](#), principalment elaborats en el món anglosaxó —tot i que molts dels estudis inclosos s'han desenvolupat en contextos més propers, com Espanya— i que es caracteritzen perquè fonamenten les aportacions en aproximacions metodològiques sòlides.

La revisió està integrada fonamentalment per nou informes que revisen en profunditat un nombre elevat d'estudis de disseny experimental o procediment metanalític i per cinc estudis experimentals que han fet aportacions especialment rellevants en el tema que ens ocupa. Bona part de les revisions inclouen, juntament amb l'evidència referida a etapes educatives obligatòries, algunes recerques i metanàlisis sobre etapes formatives superiors (universitat), sense diferenciar-ne clarament els

efectes nets a cada etapa. La inclusió en la nostra revisió dels estudis experimentals ens permet focalitzar-la en l'educació primària i secundària.

Cal tenir en consideració que la metodologia varia en els diferents estudis revisats. Si bé la major part de literatura consultada aborda la comparativa entre mecanismes d'avaluació finalista i instruments de *feedback* formatiu, en alguns casos hi ha un aprofundiment més gran que permet comparar diferents tipus de *feedback* formatiu per identificar els d'impacte més gran. A més a més, en algun cas la comparació inclou també grups de control sense avaluació ni *feedback* formatiu, amb l'objectiu de mesurar l'impacte diferencial del *feedback* formatiu i l'avaluació finalista respecte de la manca d'avaluació.

Taula 3.
Estudis revisats

Autors	Tipus de document
Klute <i>et al.</i> (2017) [7]	Revisió
Elliot <i>et al.</i> (2016) [2]	Revisió
Dunn i Mulvenon (2009) [3]	Revisió
Hattie i Timperley (2007) [6]	Revisió
Shute (2007) [5]	Revisió
Ross (2006) [8]	Revisió
Dochy <i>et al.</i> (1999) [9]	Revisió
Black i Wiliam (1998) [10]	Revisió
Kluger i DeNisi (1996) [11]	Revisió
Meusen-Beekman <i>et al.</i> (2016) [12]	Estudi experimental
De Marcos <i>et al.</i> (2010) [13]	Estudi experimental
William <i>et al.</i> (2004) [14]	Estudi experimental
McDonald i Boud (2003) [15]	Estudi experimental
Schunk (1996) [16]	Estudi experimental

Font: Elaboració pròpia

L'avaluació pot millorar els resultats educatius dels alumnes?

Els estudis sobre l'impacte de l'avaluació en el rendiment escolar dels alumnes tendeixen a identificar una millora important en els resultats tant de lectura, com de matemàtiques o de capacitat de retenció de la informació. Si bé algunes recerques conclouen que l'alumnat objecte d'avaluacions i *feedbacks* formatius de qualitat pot

arribar a incrementar el progrés acadèmic entorn de 8 mesos respecte del rendiment assolit en un curs anual, les metanàlisis més recents situen en 3 mesos el progrés addicional i en 4 mesos si el professorat ha rebut formació específica per a la implementació d'aquesta avaluació formativa [17].

Els estudis sobre l'impacte de l'avaluació en el rendiment escolar dels alumnes tendeixen a identificar una millora important en els resultats tant de lectura, com de matemàtiques o de capacitat de retenció de la informació.



Amb més o menys intensitat en funció de variables que abordarem en propers apartats, l'avaluació formativa té un impacte sobre el rendiment i els resultats escolars [7] [10]. De les àrees curriculars considerades, l'avaluació formativa mostra un major efecte en les matemàtiques, mentre que l'impacte sobre lectura i escriptura, malgrat que és observable i significatiu, és més petit [7].

L'efecte de l'autoavaluació sobre l'aprenentatge ha rebut una atenció especial en la literatura, ja que majoritàriament aquesta forma d'avaluació té una incidència positiva, malgrat que, en alguns casos, el disseny de la investigació aconsella ser curosos amb els resultats [3] [16] [8]. En general, però, les dades indiquen una millora en els resultats acadèmics dels grups que fan ús de mecanismes d'autoavaluació [15], tot i que inferiors als dels casos d'avaluació conduïda pels docents [7].

Pel que fa a l'avaluació entre companys, els resultats són també positius. Aquest tipus d'avaluació es mostra sovint més intel·ligible per als alumnes, atès que el llenguatge i l'expressió del corrector-alumne li són més propers que els del professor. A més a més, aquest tipus d'avaluació permet que l'alumne no només avalui el treball d'un company, sinó que també li aporta informació sobre el seu rendiment a través de la comparació, cosa que reverteix en el seu desenvolupament propi. En aquest cas, però, hi ha recerques que han detectat també alguns efectes negatius d'aquest tipus d'avaluació, com poden ser l'ansietat o l'estrès per la tensió derivada d'avaluar els altres però també per la comparativa amb els resultats propis [12].

Tant l'autoavaluació com l'avaluació entre companys tenen efectes diferencials en funció de la competència treballada. Així, mentre que l'autoavaluació o coavaluació es mostren efectives per al treball de les matemàtiques, en altres àrees com l'escriptura o la lectura els resultats són menys clars [7].

Bona part de la recerca recent ha parat esment en l'aportació de les noves tecnologies a l'educació (*m-learning* o aprenentatge mòbil) en general i l'avaluació, especialment a l'autoavaluació, en particular. Si bé la major part d'aquesta aportació se centra en contextos d'aprenentatge en línia, algunes recerques experimentals s'han dut a terme en entorns educatius presencials. En aquests casos, les eines digitals es configuren com a complement de l'aprenentatge presencial, sigui facilitant les proves de diagnòstic inicial, obrint possibilitats d'autoavaluació o ampliant el *feedback* entre educadors i estudiants [13]. Les dades mostren que els grups que fan ús de sistemes de noves tecnologies per a l'autoavaluació obtenen unes notes finals més bones que els grups on no hi ha avaluació, amb intensitats diferents segons l'edat i el tipus d'estudi, com abordarem més endavant.

Com ja apuntàvem, malgrat la quantitat d'investigacions, les conclusions són generals i els mecanismes de millora no són sempre clars. La major part de la literatura conclou que l'avaluació i el *feedback* formatiu comporten un increment del rendiment i l'aprenentatge dels alumnes [5] [7] [10], però també hi ha estudis que no n'identifiquen cap efecte o fins i tot n'obtenen efectes negatius [5] [11]. Aquests efectes contraris a la norma general s'expliquen des de dues lògiques. En primer lloc, per l'existència d'altres variables que interfereixen en els efectes de l'avaluació, com poden ser la qualitat docent o el nivell d'aprenentatge previ de l'alumne [14]. En aquest sentit, diverses recerques han apuntat a la interacció entre els nivells de coneixement previ i les pedagogies a l'aula, de manera que alumnes amb diversos nivells d'aprenentatge previ requereixen estratègies d'aprenentatge diferenciades i responen de forma diferent a les estructures i els continguts. Per això és important poder avaluar aquest coneixement previ com a condicionant de l'aprenentatge.

En segon terme, els errors en el disseny de l'avaluació podrien explicar-ne la manca d'efectivitat i, fins i tot, l'impacte negatiu sobre l'aprenentatge. A tall d'exemple, sembla que hi ha acord a identificar l'avaluació basada en notes i no en retorn qualitatiu com una pràctica perjudicial, però també s'observen dinàmiques no aconsellables segons moments, edats o tipus d'estudi, com abordarem a continuació [5].

Quines són les característiques de l'avaluació efectiva?

L'ampli ventall de mètodes i lògiques d'avaluació dificulta l'establiment d'un consens entorn de l'impacte general de l'avaluació formativa sobre el rendiment escolar, però alhora la gran quantitat d'instruments avaluats permet que la recerca revisada identifiqui alguns factors avaluadors que contribueixen a millorar-lo:

- **Els diferents instruments d'avaluació tenen un impacte diferencial sobre el rendiment de l'alumnat.** L'aprenentatge esdevé més elevat quan la correcció bàsica (correcte/incorrecte) s'acompanya de *feedback* facilitador, és a dir, de pistes perquè l'alumnat trobi la resposta correcta, atès que la incertesa davant el fracàs pot causar frustració i reduir la motivació per aprendre. En aquest sentit, explicar l'error concret sense fer servir frases genèriques ni remetre només a la norma general millora també notablement l'aprenentatge. En canvi, instruments com els de verificació (correcte/incorrecte) o la repetició de l'exercici fins que sigui correcte plantegen impactes més reduïts o nuls. Els *feedbacks* basats en premis o càstigs tenen poc efecte. Sembla que els efectes varien en funció del tipus d'aprenentatge: si es tracta de millorar el mateix exercici, els mecanismes de repetició són més efectius; l'efecte es trasllada a la millora de tasques diferents quan els instruments d'avaluació incorporen un acompanyament més gran [5] [6].
- **El *feedback* autoreferencial és més efectiu, atès que ubica l'alumne en el seu propi procés d'aprenentatge.** L'assignació de notes, com la resta de mecanismes de *feedback* normatiu, esdevé un mecanisme d'impacte inferior i comporta, sovint, efectes negatius. En el



cas concret de l'assignació de notes, s'observa una reducció de l'impacte de l'avaluació, atès que centra la preocupació de l'alumnat en la puntuació i no en els comentaris formatius que acompanyen la correcció [2] [5].

- **El *feedback* ha d'adequar-se a l'alumnat a qui es dirigeix.** Quan no hi ha coherència entre missatge i receptor, no s'observa impacte del *feedback* sobre el rendiment de l'alumnat. A tall d'exemple, l'alumnat de nivell elevat es beneficia de nivells més baixos de detall en el *feedback*, mentre que el més motivat obté un rendiment més gran a partir de mecanismes més complexos d'avaluació. Així mateix, els alumnes de baix rendiment es beneficien més del retorn o la correcció immediata, mentre que sembla que els d'alt rendiment es beneficien més del retorn en diferit [5].

- **La recerca no és concloent respecte al grau d'elaboració i complexitat aconsellable per fer retorn de l'avaluació.**

La literatura revisada conclou que qualsevol retorn avaluador té més efecte que la simple correcció i que la complexitat del mecanisme emprat té un efecte diferencial en funció d'allò que s'avalua. No obstant això, no són clares les conclusions respecte a quin nivell de complexitat comporta guanys més grans i alguns estudis determinen que no hi ha cap efecte diferencial segons el grau de complexitat [5].

La literatura revisada conclou que qualsevol retorn avaluador té més efecte que la simple correcció i que la complexitat del mecanisme emprat té un efecte diferencial en funció d'allò que s'avalua.



- **El *feedback* té un impacte més gran quan els objectius de l'aprenentatge són molt clars.** L'avaluació es mostra així com un complement de les dinàmiques formatives desenvolupades a l'aula, i n'emfasitza els efectes quan tota l'activitat educativa s'organitza d'acord a uns objectius establerts [11].

- **Els efectes sobre el rendiment escolar de l'avaluació són més elevats quan la complexitat de la tasca és baixa.** A mesura que l'activitat guanya en complexitat, es redueix l'aportació del retorn formatiu al rendiment escolar [11].

El *feedback* té un impacte més gran quan els objectius de l'aprenentatge són molt clars. L'avaluació es mostra així com un complement de les dinàmiques formatives desenvolupades a l'aula.



- **L'avaluació orientada a l'aprenentatge que no es limita a la valoració del rendiment en cada prova planteja un impacte més gran.** L'avaluació i el *feedback* permeten una millora més significativa del rendiment quan s'enfoquen a l'aprenentatge en conjunt i no quan es limiten a la correcció del rendiment en cada prova. En aquest darrer cas, l'impacte pot ser negatiu o inconsistent [5].

- **L'avaluació i el *feedback* poden ser immediats o en diferit.** Algunes recerques identifiquen que el *feedback* immediat evita la memorització d'errors, mentre que d'altres consideren que pot interferir abans de temps, sense deixar a l'alumnat espai per processar i corregir sol l'error. Vinculat a això, diversos estudis aconsellen variar el moment de correcció en funció del perfil de l'alumnat: mentre que el de baix rendiment es beneficia més de la correcció immediata, el més avançat millora més clarament quan el *feedback* es fa de forma diferida [5].

- **Hi ha consens a qualificar l'avaluació formativa que es produeix de forma contínua durant el procés d'aprenentatge com l'avaluació de més impacte.** Si bé

l'avaluació final i la diagnòstica són també rellevants, diversos estudis sobre l'avaluació del progrés o avaluació contínua indiquen una reducció important de l'impacte de factors individuals sobre el rendiment a partir de la realització de proves de progrés amb el corresponent retorn al llarg del curs [9].

- Malgrat que tant l'autoavaluació com la coavaluació i l'avaluació dirigida per docents són efectives, **l'impacte sobre el rendiment escolar és més gran en els casos que el professorat és qui condueix i monitora el progrés de l'alumnat [7].**

Malgrat que tant l'autoavaluació com la coavaluació i l'avaluació dirigida per docents són efectives, l'impacte sobre el rendiment escolar és més gran en els casos que el professorat és qui condueix i monitora el progrés de l'alumnat.



Requadre 1.

Formació en autoavaluació i impactes educatius. Experiment a Barbados

Betty McDonald i David Boud [13] van dur a terme un experiment a deu centres d'educació secundària a l'illa de Barbados. Amb l'objectiu d'avaluar l'impacte de les habilitats d'autoavaluació sobre el rendiment escolar dels alumnes, van dissenyar un programa de formació en autoavaluació per a alumnes i professors.

Es van seleccionar deu escoles segons els resultats obtinguts a les proves d'avaluació externa realitzades anualment pel Caribbean Examinations Council i les proves d'accés a l'educació secundària que tot l'alumnat ha de passar per tal d'escollir centre (Barbados Secondary Schools Entrance Examination). Aquests deu instituts responen a una mostra d'escoles d'alt, mitjà i baix rendiment segons aquests resultats.

A cada centre es van seleccionar dos grups classe de l'últim curs de l'educació secundària (equivalent a 4t d'ESO), de 30 a 40 alumnes. Els investigadors van revisar, centre per centre, les avaluacions dels cursos anteriors dels alumnes de les dues classes per tal d'assegurar que no es produïa una agrupació per nivells que pogués esbiaixar la mostra. A cada escola es va aleatoritzar quina de les dues classes quedaria assignada al grup de tractament i quina al grup de control. Així, un total de 256 alumnes van rebre formació en habilitats d'autoavaluació, mentre que el grup de control (259 alumnes) no va rebre aquesta formació.

El grup experimental va rebre formació en autoavaluació durant el curs escolar i com a part del currículum escolar ordinari a partir de 12 mòduls dissenyats pels investigadors. De la mateixa manera, els docents van ser instruïts per tal de facilitar a l'alumnat aquesta formació. El grup de control va seguir el currículum ordinari sense incloure aquesta formació.

Al final de curs, tots els alumnes van fer les proves del Caribbean Examinations Council (CxC). Els resultats entre ambdós grups van diferir significativament. El grup tractat (amb formació en autoavaluació) hi va obtenir unes mitjanes superiors en totes les àrees curriculars avaluades (economia, humanitats, ciències i tecnologia).

Els investigadors indiquen que cal tenir en consideració els efectes no buscats que la formació en autoavaluació ha pogut tenir sobre altres variables també explicatives del rendiment escolar, com pot ser la motivació dels docents implicats en el procés o la motivació per part de l'alumnat del grup de tractament.

Per a més informació:

McDonald, B. i Boud, D. (2003). "The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations". *Assessment in Education. Principles, Policy and Practice*, vol. 10, núm. 2, p. 209—220. [15]

Quin alumnat és més sensible a l'avaluació formativa?

Més que indicar quin alumnat és més sensible a l'avaluació formativa, cal preguntar-se quins perfils es beneficien més dels diferents instruments d'avaluació formativa. En altres paraules, quins són els efectes d'aquests instruments sobre els diferents tipus d'alumnat:

- L'alumnat d'alt rendiment obté resultats més bons amb els mecanismes menys intensos (com la simple verificació correcte/error). En canvi, el de baix rendiment experimenta un aprenentatge més gran amb mecanismes més intensius de *feedback* facilitador, que incloquin l'explicació detallada de l'error i la possibilitat de correcció per part de l'alumne. Per la seva banda, els alumnes de nivell mitjà no mostren grans diferències de rendiment en funció del mecanisme de *feedback* emprat. En qualsevol cas, els tres perfils d'alumnes mostren millores substancials en comparació d'aquells que no han rebut cap retorn [5].

L'alumnat d'alt rendiment obté resultats més bons amb els mecanismes menys intensos (com la simple verificació correcte/error). En canvi, el de baix rendiment experimenta un aprenentatge més gran amb mecanismes més intensius de *feedback* facilitador.



- L'alumnat de baix rendiment es beneficia més del retorn o correcció immediata, mentre que el d'alt rendiment es beneficia més del retorn en diferit [5].
- L'alumnat de baix rendiment és més sensible al tipus de retorn que rep. Quan el *feedback* és normatiu (comparació respecte al conjunt de la classe), els alumnes de perfil baix tendeixen a atribuir els seus resultats a una manca d'habilitat, i no esperen millorar-los en proves següents, cosa que n'incrementa la desmotivació. En canvi, el *feedback* autoreferencial (comparació respecte a les capacitats de l'alumne mateix) vincula directament els resultats a l'esforç que ha fet, cosa que fa que l'alumnat focalitzi l'atenció en el seu progrés particular [5].

L'alumnat de baix rendiment és més sensible al tipus de retorn que rep. Quan el *feedback* és normatiu, els alumnes de perfil baix tendeixen a atribuir els seus resultats a una manca d'habilitat, i no esperen millorar-los en proves següents, cosa que n'incrementa la desmotivació.



- Les recerques centrades en alumnat amb discapacitat mostren conclusions més clares que en el cas de l'alumnat sense discapacitat. Així doncs, malgrat la falta d'acord entre algunes recerques respecte de l'impacte diferencial en comparació amb el conjunt de la població, s'observa unanimitat entre aquells estudis que indiquen que l'alumnat d'educació especial es beneficia d'una manera intensa de la introducció de mètodes avaluadors formatius [3] [10].
- El nivell socioeconòmic de l'alumnat ha estat poc abordat per la literatura revisada, però els estudis centrats en escoles en entorns socials més empobrits identifiquen impactes per sobre del conjunt dels estudis, fet que podria indicar una millora en el rendiment d'aquest alumnat superior a la del conjunt [3] [10]. Es confirma així que l'alumnat empobrit és clarament sensible als efectes de l'avaluació.
- Algunes recerques suggereixen que el *feedback* formatiu només té impacte (o en qualsevol cas en té més) en l'alumnat més jove. Tot i això, la identificació d'aquestes

edats varia en funció de la recerca. Hi ha consens, en canvi, a indicar que l'avaluació formativa requereix un aprenentatge, de manera que com més aviat s'iniciï, més gran en serà l'impacte. En el cas de l'avaluació basada en mètodes digitals, també es remarca la familiaritat més gran amb les noves tecnologies entre els més joves com un factor de motivació que n'incrementa l'impacte [3] [13].

A més del rendiment, quines altres competències pot millorar l'avaluació?

La recerca sobre avaluació i *feedback* formatiu centra l'atenció en l'aportació d'aquests instruments a la millora de l'aprenentatge (present i futur), i no només dels resultats escolars. En aquest sentit, hi ha tres aspectes freqüentment abordats per les revisions i els estudis analitzats: la motivació per l'aprenentatge, la capacitat d'autoregulació o autoorganització de l'aprenentatge per part de l'alumnat i la seva percepció d'autoeficàcia [10] [15]. Avancem que la literatura observa un impacte positiu de l'avaluació formativa només en els dos primers.

- **Motivació per l'aprenentatge.** Les actituds positives envers l'aprenentatge s'incrementen a partir de l'avaluació formativa. La confiança generada per la capacitat d'avaluació de l'alumnat reverteix en un increment de la motivació per aprendre. Cal indicar, també, un increment de la motivació per l'aprenentatge a través de mecanismes d'aprenentatge mòbil o *m-learning* en els grups d'edat més jove, segurament pel fet que estan més familiaritzats amb les noves tecnologies, mentre que en estudiants d'edats més avançades no s'observen variacions [13].

Les actituds positives envers l'aprenentatge s'incrementen a partir de l'avaluació formativa. La confiança generada per la capacitat d'avaluació de l'alumnat reverteix en un increment de la motivació per aprendre.



- **Capacitat d'autoorganització de l'aprenentatge.** L'autoregulació de l'aprenentatge no és una tasca senzilla i requereix moltes habilitats, des de la capacitat de planificació i establiment d'objectius fins a l'aptitud per reflexionar i posar en marxa les estratègies adequades per a l'aprenentatge.² Valorada tradicionalment en etapes educatives superiors (educació secundària, universitat, etc.), la capacitat d'autogestió i organització de l'aprenentatge per part de l'alumnat és, cada cop més, un objectiu també de l'etapa primària. La cerca d'estratègies que els ajudin a organitzar-se les feines durant el curs escolar ha esdevingut part de la tasca del professorat d'educació primària i, entre aquestes estratègies, l'avaluació formativa es consolida com una eina. Tot i que no hi ha una edat per desenvolupar la capacitat d'autogestió, la recerca posa de manifest que, amb metodologies adaptades a etapes primerenques, les estratègies d'autoregulació de l'aprenentatge durant l'educació primària tenen efectes positius en etapes posteriors. I l'avaluació

La cerca d'estratègies que els ajudin a organitzar-se les feines durant el curs escolar ha esdevingut part de la tasca del professorat d'educació primària i, entre aquestes estratègies, l'avaluació formativa es consolida com una eina.



² Per a més detalls sobre el concepte d'autoregulació i sobre l'efectivitat de les intervencions educatives que miren de fomentar-la, podeu consultar l'article de Gerard Ferrer-Esteban en el número 5 de la sèrie "Què funciona en educació": <https://goo.gl/s6WC1Y>.

formativa sembla que contribueix al desenvolupament d'aquestes habilitats, especialment quan aquesta perspectiva s'inclou en totes les activitats a l'aula [12]. L'aportació de l'avaluació formativa rau en la informació sobre la distància entre el rendiment de l'alumne i l'esperat, però també i sobretot en les eines que atorga per a la millora del rendiment, tot assenyalant quines són les millors estratègies a seguir. És en aquest punt que l'autoregulació es veu incrementada, atès que l'alumnat rep un ventall ampli d'opcions per incorporar i desenvolupar en el futur i, per tant, se'ls permet ser responsables del seu aprenentatge [14]. L'impacte és més gran sobre l'autoregulació quan l'autoavaluació és el mecanisme avaluador, tot i que l'avaluació entre companys també incrementa la capacitat d'autoregulació de l'alumnat [12].

- **Percepció d'autoeficàcia.** La percepció de l'alumnat respecte de la seva pròpia capacitat per executar tasques o assolir els objectius fixats és un element rellevant del procés d'aprenentatge. Malgrat que es tracta d'un aspecte molt proper i sovint relacionat amb la capacitat d'autoregulació, la recerca no observa efectes significatius entre el mètode avaluador i la percepció de l'alumne respecte de les seves capacitats [12].

Resum

Del conjunt de lògiques i instruments avaluadors, en aquesta revisió hem posat l'atenció només en aquells amb una voluntat formativa explícita, tot i que en algun moment els estudis han fet servir en el grup de control tècniques avaluadores finalistes, no formatives. La revisió de l'evidència ha mostrat uns impactes significatius del *feedback* formatiu sobre el rendiment acadèmic i també sobre qüestions més àmplies d'aprenentatge, com la capacitat d'autoorganització o la motivació per aprendre. En canvi, sembla que les percepcions sobre les capacitats pròpies no són sensibles als mètodes avaluadors.

L'amplitud i varietat d'instruments avaluadors dificulta obtenir resultats generalitzables, però ahora permet la identificació d'algunes pràctiques d'impacte més gran. S'observa la simple verificació (correcte/incorrecte) o les dinàmiques de premis i càstigs com a instruments de poc impacte sobre el rendiment de l'alumnat. En canvi, l'acompanyament de la correcció amb un *feedback* facilitador, l'autoavaluació per part dels alumnes o l'avaluació entre companys es configuren com a opcions d'interès per a la millora del rendiment.

L'acompanyament de la correcció amb un *feedback* facilitador, l'autoavaluació per part dels alumnes o l'avaluació entre companys es configuren com a opcions d'interès per a la millora del rendiment.



Cal destacar, a més a més, que som en un àmbit amb impactes clarament diferencials segons el perfil de l'alumnat. Així, mentre que l'alumnat de baix rendiment es beneficia de sistemes de retorn formatiu més complexos i més immediats, l'alumnat més avançat obté uns resultats més bons amb *feedbacks* més simples i realitzats en diferit. Aquests dos perfils també es mostren diferencialment sensibles als tipus d'orientació del *feedback*. En general, l'alumnat més vulnerable —sigui per capacitats, sigui per

classe social— sembla mostrar una millora més clara del rendiment escolar a partir de l'aplicació d'instruments d'avaluació formativa, especialment quan s'adapta a les seves característiques.

Taula 4.

Arguments a favor i limitacions dels instruments d'avaluació formativa

Punts a favor	Limitacions
<ul style="list-style-type: none"> • L'avaluació formativa té uns efectes més grans sobre el rendiment que l'avaluació finalista. 	<ul style="list-style-type: none"> • L'avaluació ha de considerar la vocació de formació i no limitar-se a la valoració dels aprenentatges per tal de tenir efectes sobre el rendiment.
<ul style="list-style-type: none"> • L'avaluació facilitadora contribueix més a l'aprenentatge que l'avaluació directiva. 	<ul style="list-style-type: none"> • L'avaluació ha d'evitar la comparació respecte als altres alumnes i centrar la valoració respecte al potencial de l'alumne en qüestió.
<ul style="list-style-type: none"> • Cal establir una connexió clara entre els criteris d'avaluació i els objectius de les unitats d'aprenentatge per fer de l'avaluació un instrument formatiu. 	<ul style="list-style-type: none"> • L'assignació de notes redueix el potencial efecte formatiu de l'avaluació.
<ul style="list-style-type: none"> • Cal acompanyar la identificació dels errors amb pistes per resoldre'ls o millorar-los. 	<ul style="list-style-type: none"> • S'ha d'evitar el retorn desmotivador centrat en els errors; cal, en canvi, emfasitzar les recomanacions o suports per millorar.
<ul style="list-style-type: none"> • L'avaluació ha d'adreçar-se a l'aprenentatge en conjunt. 	<ul style="list-style-type: none"> • L'avaluació formativa no pot limitar-se a la valoració del rendiment en una prova concreta.
<ul style="list-style-type: none"> • El <i>feedback</i> s'ha d'organitzar en unitats comprensibles per a l'alumne. 	<ul style="list-style-type: none"> • L'avaluació no pot ser complexa, atès que s'ha d'assegurar que s'entén i que dota l'alumne d'eines per resoldre errors previs.
<ul style="list-style-type: none"> • Per incentivar un aprenentatge general, no limitat a la qüestió avaluada, el retorn en diferit mostra efectes més bons. 	<ul style="list-style-type: none"> • El retorn en diferit no s'aconsella per a tasques molt complexes ni en el cas d'alumnat amb dificultats d'aprenentatge.

Font: Elaboració pròpia

Implicacions per a la pràctica

Les dades mostrades en aquest article posen de manifest l'important impacte de l'avaluació formativa sobre el rendiment de l'alumnat, però alhora també la seva fragilitat, atesa la multitud de factors que poden reduir o emfasitzar aquests efectes. En altres paraules, els beneficis de l'avaluació no estan garantits sota qualsevol circumstància, sinó que calen diferents requisits perquè es donin, alguns dels quals depenen dels responsables polítics. En aquest sentit s'exposen a continuació algunes recomanacions:

- És important que els esforços d'avaluació per part del personal docent es facin de la forma més eficient i eficaç possible. Per això, cal formar el professorat en instruments d'avaluació formativa i acompanyar les escoles en el disseny dels plans d'avaluació, amb l'objectiu d'incrementar el rendiment de l'alumnat.

- El context s'ha mostrat molt rellevant per a l'èxit de l'avaluació formativa. Aquesta s'ha de plantejar en el marc d'un procés d'aprenentatge no finalista, és a dir, orientat a l'aprenentatge continu. Bona part de l'èxit de l'avaluació deriva de la identificació d'objectius clars per a l'alumnat, de manera que, juntament amb els criteris de correcció, cal clarificar els objectius de l'aprenentatge per tal que facin entenedor el sentit de les tasques i les avaluacions. És un procés llarg, que ha d'iniciar-se amb la planificació de l'activitat i ha de continuar fins a l'avaluació i el retorn. Així doncs, cal dotar de temps el personal docent perquè planifiqui l'avaluació com a part del procés d'aprenentatge.
Cal dotar de temps el personal docent perquè planifiqui l'avaluació com a part del procés d'aprenentatge.

- L'impuls de l'avaluació permanent ha de permetre ampliar els objectius més enllà del rendiment, cap a altres esferes de l'aprenentatge. En aquest sentit, el treball d'autoavaluació o avaluació entre companys a l'educació primària sembla tenir efectes importants sobre la capacitat d'autoregulació de l'alumnat i pot comportar millores en la transició a la secundària, etapa en què el control del docent disminueix i es requereixen més capacitats d'autogestió. L'autoavaluació o l'avaluació entre companys requereix una planificació especial, recursos per desenvolupar-la i la formació dels docents per posar-la en marxa.
- D'altra banda, si es considera el pes explicatiu que els coneixements previs tenen sobre el rendiment dels alumnes i la certesa que aquests coneixements condicionen l'aprenentatge futur, sembla oportú dotar els centres escolars d'eines per a l'avaluació diagnòstica que permetin afinar el disseny del procés d'aprenentatge. En aquest sentit, calen recursos i formació per identificar adequadament els coneixements previs. Així mateix, seria convenient que l'administració pública fes servir aquesta avaluació inicial com a mecanisme de control en l'anàlisi de l'avaluació de docents i centres educatius.
Sembla oportú dotar els centres escolars d'eines per a l'avaluació diagnòstica que permetin afinar el disseny del procés d'aprenentatge.

- L'administració pública necessita dades i informació per millorar el disseny dels processos educatius. Amb aquest objectiu, les avaluacions esdevenen de gran importància. Cal, però, que aquestes avaluacions siguin completes i permetin captar el nombre més gran de variables que intervenen en el procés (capital instructiu familiar, motivació, qualitat docent...). Cal, a més a més, que aquestes avaluacions estiguin a disposició del personal docent i de l'alumnat perquè les emprin amb finalitat formativa.
- Per acabar, en vista de l'ambigüitat d'alguns resultats, cal impulsar recerca que avaluï més en profunditat l'impacte dels diferents instruments avaluadors, la intermediació d'altres variables com la qualitat del professorat i, molt especialment, els efectes respecte de l'alumnat de baix rendiment.

Bibliografia

- [1] OCDE (2013). *Synergies for Better Learning*. OECD Publishing.
- [2] Elliot, V.; Baird, J.-A.; Hopfenbeck, T.; Ingram, J.; Thompson, I.; Usher, N.; Zantout, M.; Richardson, J. i Coleman, R. (2016). "A marked improvement?".
- [3] Dunn, K. E. i Mulvenon, S. W. (2009). "A Critical Review of Research on Formative Assessment: The Limited Scientific Evidence of the Impact of Formative Assessment in Education". *Pract. Assessment, Res. Eval.*, vol. 14, núm. 7, p. 1-11.
- [4] Bell, B. i Cowie, B. (2000). "The Characteristics of Formative Assessment in Science Education". *Sci. Educ.*, vol. 85, p. 536-553.
- [5] Shute, V. J. (2007). *Focus on Formative Feedback*, núm. de març.
- [6] Hattie, J. i Timperley, H. (2007). "The Power of Feedback". *Rev. Educ. Res.*, vol. 77, núm. 1, p. 81-112.
- [7] Klute, M.; Apthorp, H.; Harlacher, J. i Reale, M. (2017). "Formative assessment and elementary school student academic achievement: A review of the evidence".
- [8] Ross, J. A. (2006). "The Reliability, Validity, and Utility of Self-Assessment". *Pract. Assessment, Res. Eval.*, vol. 11, núm. 10, p. 1-11.
- [9] Dochy, F.; Segers, M. i Buehl, M. M. (1999). "The Relation Between Assessment Practices and Outcomes of Studies: The Case of Research on Prior Knowledge". *Rev.*, vol. 69, núm. 2, p. 145-186.
- [10] Black, P. i Wiliam, D. (1998). "Assessment and Classroom Learning". *Assess. Educ. Princ. Policy Pract.*, vol. 5, núm. 1, p. 7-74.
- [11] Kluger, A. N. i DeNisi, A. (1996). "The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory". *Psychol. Bull.*, vol. 119, núm. 2, p. 254-284.
- [12] Meusen-Beekman, K. D.; Brinke, D. J. i Boshuizen, H. P. A. (2016). "Studies in Educational Evaluation Effects of formative assessments to develop self-regulation among sixth grade students: Results from a randomized controlled intervention". *Stud. Educ. Eval.*, vol. 51, p. 126-136.
- [13] De Marcos, L.; Hilera, J. R.; Barchino, R.; Jiménez, L.; Martínez, J. J.; Gutiérrez, J. A.; Gutiérrez, J. M. i Otón, S. (2010). "An experiment for improving students performance in secondary and tertiary education by means of m-learning auto-assessment". *Comput. Educ.*, vol. 55, núm. 3, p. 1069-1079.
- [14] Wiliam, D.; Lee, C.; Harrison, C. i Black, P. (2004). "Teachers developing assessment for learning: impact on student achievement". *Assess. Educ. Princ. Policy Pract.*, vol. 11, núm. 1, p. 49-65.
- [15] Mcdonald, B. i Boud, D. (2003). "The Impact of Self-assessment on Achievement: The effects of self-assessment training on performance in external examinations". *Assess. Educ. Princ. Policy Pract.*, vol. 10, núm. 2, p. 209-220.
- [16] Schunk, D. H. (1996). "Goal and Self-Evaluative Influences During Children's Cognitive Skill Learning". *Am. Educ. Res. J.*, vol. 33, núm. 2, p. 359-382.
- [17] Education Endowment Foundation Foundation (2017). "Teaching & Learning Toolkit: Feedback".

Primera edició: maig de 2018
© Fundació Jaume Bofill, Ivàlua, 2018
fbofill@fbofill.cat, info@ivalua.cat
www.ivalua.cat
www.fbofill.cat

Autora: Sheila González Motos
Edició: Bonalletra Alcompàs
Coordinació editorial: Anna Sadurní
Cap de projectes internacionals: Valtencir Mendes
Disseny i maquetació: Enric Jardí
ISBN: 978-84-947888-4-0

Aquesta obra està subjecta a la llicència Creative Commons de **Reconeixement-No Comercial-SenseObraDerivada (by-nc-nd)**. Es permet la reproducció, distribució i comunicació pública de l'obra sempre que se'n reconegui l'autoria. No es permet l'ús comercial de l'obra ni la generació d'obres derivades.

