



Çfarë është vlerësimi për të nxënëit?

Dylan Wiliam

Instituti i Arsimit, Universiteti i Londrës, Mbretëri e Bashkuar

INFO ARTIKULLI

Historia e artikullit:
Në dispozicion online 12 prill 2011

Fjalë kyçe:
Vlerësim formues
Vlerësimi për të nxënëit
Reagimi

ABSTRAKT

Ideja se vlerësimi është thelbësisht i domosdoshëm për një mësimdhënie efektive e ka origjinën në eksperimentet e hershme mbi mësimin individual që nga vepra e Benjamin Bloom deri te rishikimet mbi ndikimin që reagimi ka mbi nxënësit në klasë. Ndërkohë që shumë prej rishikimeve paraqesin në hollësi ndikimin negativ të vlerësimit në mësim, ato gjithashtu sugjerojnë se në kushte të caktuara, vlerësimi mund të ketë mundësi të konsiderueshme për të përmirësuar mësimin. Kjo tregon se të kuptuarit e ndikimit që vlerësimi ka mbi mësimin kërkon një fokus më të gjerë sesa vetë reagimi, veçanërisht reagimet e nxënësit ndaj reagimit, dhe mjedisi mësimor në të cilin reagimi funksionon. Përkufizimet e ndryshme të termave “vlerësim formues” dhe “vlerësim për mësim” janë trajtuar dhe përfshirë nën një përkufizim të gjerë që përqendrohet në masën në të cilën vendimet udhëzuese mbështeten në prova. Punimi përfundon duke shqyrtuar disa nga pasojat e këtij përkufizimi për praktikën në klasë.

© 2011 Elsevier Ltd. Të gjitha të drejtat e rezervuara.

1. Hyrje

Rreth një gjysmë shekulli më parë, David Ausubel sugjeroi se faktori më i rëndësishëm që ndikon në mësim është çfarë njohurish ka tashmë nxënësi, mësuesit konstatojnë këtë fakt, dhe japin mësim në përputhje me rrethanat. (Ausubel, 1968).

Në atë kohë, dhe mbase edhe në ditët e sotme, një përshkrim i tillë mund të duket i thjeshtë, por nxënësit nuk mësojnë çfarë i shpjegohet. Edhe kur shpjegimi është planifikuar me shumë kujdes, dhënë në mënyrë efektive dhe tërheqëse për nxënësit, rezultatet e mësimit shpesh kanë pak lidhje ose aspak lidhje me synimin. Nëse ajo që nxënësi mëson si rezultat i një rendi të caktuar të aktiviteteve mësimore është e pamundur të parashikohet, edhe në rastin e vështirë ku të gjithë nxënësit e një grupi mësimor janë në të njëjtin nivel në fillim të mësimit, brenda disa minutave, nxënësit marrin kuptime të ndryshme. Kjo mbase është arsyeja përse vlerësimi është procesi kryesor për një mësim efektiv. Vetëm nëpërmjet vlerësimit mund të kuptojmë nëse një rend i caktuar e aktiviteteve mësimore ka rezultuar i suksesshëm në rezultatet e synuara të mësimit.

Për shumë vite, fjala “vlerësim” është përdorur kryesisht për të përshkruar proceset e vlerësimit të efektivitetit të kursit të aktiviteteve mësimore kur ka përfunduar. Veprimet që udhëzuan proceset mësimore përpara përfundimit të kursit, në përgjithësi nuk janë konsideruar si lloje vlerësimesh. Në letërsinë franceze, janë konsideruar zakonisht si aspekte të rregullimit të proceseve mësimore, dhe në literaturën angleze, thjesht si një aspekt i një mësimdhënie të mire.

Megjithatë së fundmi, veçanërisht në komunitetin kërkimor angliisht-folës,

0191-491X/\$ - see front matter © 2011 Elsevier Ltd. Të gjitha të drejtat e rezervuara. doi:10.1016/j.stueduc.2011.03.001

një tendencë në rritje për të kuptuar aktivitetet që kanë për qëllim të udhëheqin mësimin drejt qëllimit të synuar dhe kjo ndodh gjatë procesit mësimor, si forma vlerësimi. Ky punim përshkruan këtë zhvillim dhe përpjet të qartësojë kuptimet e termave *vlerësim për mësim* dhe *vlerësim formues*.

2. Vlerësimi për mësimin: Origjina dhe pararendësit

Për shumë vite, duket se pikëpamja mbizotëruese në arsim ishte se nëse mësimi gjykohej si cilësor, nuk ishte e nevojshme të përshtatej me nevojat e nxënësve. Konsiderohej se mësimi i mirë-përcaktuar do të ishte efektiv për shumicën e nxënësve për të cilët ishte menduar (ndërsa të tjerët do të ishin pjesë e aktiviteteve përmirësuese) apo se vet nxënësit ishin shkaku i dështimit të mësimit (materiali ishte shumë i vështirë për ta, dhe se duhet të ndjekin një material tjetër, dhe në përgjithësi rrugë më pak akademike). Pavarësisht kësaj, në vitet '60, Benjamin Bloom dhe nxënësit e tij në Universitetin e Çikagos filluan të eksploronin idenë se shpërndarja normale e rezultateve të nxënësve nuk ishte një rezultat “i natyrshëm”, por i shkaktuar nga mosarritja e bërjes së dallimeve ndërmjet nxënësve.

“Sistemi Individual” shpesh vlerësuar si sistemi i parë i individualizuar i mësimdhënies (shih, për shembull, Reiser, 1986), u krijua nga Frederic Burk, nga viti 1912 deri më 1913, për t'u përdorur në shkollën fillore të lidhur me Shkollën Shtetërore të San Françiskos, një institucion që ofron shërbime trajnimi për mësuesit. Një nga kolegët e Burk, Mary Ward, merrej me mësuesit në trajnim për të zhvilluar materiale vetë-mësimore dhe

Burk dhe të tjerë hartuan materiale të ngjashme që mbulonin shumicën e kurrikulës nga arsimi parashkollor deri në klasën e gjashtë. Dy persona të tjerë të cilët kishin punuar me Ward dhe Burk në shkollën shtetërore të San Françiskos, Carleton Washburne dhe Helen Parkhurst, zhvilluan këto ide akoma më tepër pas largimit të tyre nga shkolla. Në vitin 1919, Washburne vuri në zbatim Planin Winnetka, në kohën kur ishte mbikëqyrës i shkollave publike në Winnetka në Illinois dhe në të njëjtin vit, Parkhurst vuri në zbatim Planin Dalton në një shkollë për nxënësë me aftësi të kufizuara në Dalton, Massachusetts (Parkhurst, 1922).

Bloom ishte i bindur se një individualizim i tillë ishte i dobishëm—ai e konsideronte mësimin individual si “standardin e artë” për edukim përkundrejt krahasimeve me të tjerë (Bloom, 1984)—por ishte i shqetësuar që kjo padyshim nuk do ishte e përballeshme për arsimin publik në masë, dhe për rrjedhojë “kërkimi për metoda mësimi në grup po aq të efektshme sa mësimi individual” (Bloom, 1984b).

Sipas Bloom, një nga arsyt kryesore përse mësimi individual është efektiv është se mësuesi është në gjendje të identifikojë gabimet e nxënësit menjëherë, mund të japë sqarime mbi to si dhe ta ndjekë në vazhdimësi nëse është e domosdoshme (Guskey, 2010). Bloom i përshkruante këto dy procese si “reagim” dhe “korrigjime” dhe që nga ajo kohë kjo gjuhë është bërë pjesë e mënyrës së folurit mbi vlerësimin. Megjithatë, me shumë rëndësi është fakti se dallimi që Bloom bën ndërmjet “reagim” dhe “korrigjimeve” ka qenë jo produktiv, dhe fatkeqësisht ka shërbyer për të deformuar kuptimin origjinal të termit “reagim”.

Në vitin 1940, Norbert Wiener dhe kolegët e tij ishin duke ndërtuar gjetje automatike për armë kundër raketore. Ai kuptoi se një veprim efektiv kishte nevojë për një sistem të mbyllur që lejonte efektet e veprimeve të marra brenda sistemit për t’u vlerësuar, dhe në bazë të këtij vlerësimi, të modifikohen veprimet në të ardhmen (Wiener, 1948). Në sisteme të tilla, ekzistojnë dy lloj hapësirash: ata që priren të shtyjnë sistemin në drejtimin në të cilin po shkonte (të cilat i quajti hapësira pozitive reagimi) dhe ato që shkonin kundër tendencës së sistemit (të cilat i quante hapësira negative reagimi). Hapësirat pozitive të reagimit paraqesin paqëndrueshmëri, duke e çuar sistemin drejt eksplozionit apo kollapsit. Shembuj janë rritja e popullsisë me ushqim të bollshëm dhe pa grabitqarë, dhe vala e inflacionit në çmime/paga në ekonomi; shembuj të së dytës përfshijnë depresionin ekonomik, ruajtja e ushqimit në kohë mungese, dhe humbja e të ardhurave tatimore në zonat urbane si rezultat i “fluturimit të klasës së mesme”. Përdorimi i kualifikimit “pozitiv” nuk ka për qëllim të sigurojë ndonjë vlerësim të vlerës së një reagimi të tillë – në të vërtetë reagimi pozitiv pothuajse gjithmonë ka efekte të padëshirueshme. Përkundrazi, termi “pozitiv” nënkupton thjesht përafrimin ndërmjet tendencës ekzistuese të sistemit dhe efektin e impulsit të dhënë nga reagimi.

Nga ana tjetër, hapësirat negative të reagimit sigurojnë qëndrueshmëri, pasi kanë prirje të drejtojnë sistemin në gjendjen e mëparshme. Shembull i një sistemi të tillë është rritja e popullsisë me furnizim të kufizuar ushqimor, ku mungesa e ushqimit ngadalëson rritjen e popullsisë. Në varësi të kushteve, sistemi ose afrohet ose luhet me amplitudë në rënie, një gjendje e qëndrueshme (kapaciteti i mjedisit). Ndoshta shembulli më i njohur është termostati shtëpiak. Kur temperatura e dhomës bie nën nivelin e termostatit, dërgohet një sinjal për të ndezur sistemin e ngrohjes. Kur dhoma nxehet mbi nivelin e termostatit, dërgohet një sinjal për të fikur sistemin e ngrohjes.

Një pikë e rëndësishme në formulimin e Wiener është se

informacioni nuk behët “reagim” përveç se nëse jepet brenda një sistemi që mund ta përdorë atë informacion për të ndikuar në të ardhmen. Rëndësia e të menduarit mbi sistemet e reagimit, në vend të vet natyrës së informacionit, veçanërisht në shkencat e sjelljes, u theksua nga Ramaprasad

(1983) i cili vuri në dukje se: “Reagimi [feedback] është informacioni rreth hapësirës ndërmjet nivelit aktual dhe nivelit referues së një sistemi, parametër që përdoret në një farë mënyre për të ndryshuar hapësirën” (Ramaprasad, 1983, f. 4). Përdorimi i informacionit u përforcua nga Sadler (1989):

Një tipar i rëndësishëm i përcaktimit të Ramaprasad është se informacioni rreth hapësirës ndërmjet niveleve aktuale dhe referuese konsiderohet si reagim vetëm kur përdoret për të ndryshuar hapësirën. Nëse informacioni thjesht regjistrohet, i kalohet një pale të tretë e cila nuk ka njohuri apo fuqi për të ndryshuar rezultatin ose është tepër i kodifikuar (për shembull, si notë përmbledhëse dhënë nga mësuesi) për të ndërmarrë veprimin e duhur, hapësira e kontrollit nuk mund të mbyllet, dhe “të dhënat” zëvendësohen me reagimin efektiv. (f. 121)

Kjo është arsyeja përse formulimi i Bloom nuk është i dobishëm. Në përshkrimin e informacionit të dhënë mbi hapësirën ndërmjet performancës aktuale dhe asaj të dëshiruar si “reagim” Bloom veçonte informacionin nga pasojat e tij mësimore. Për Wiener, Ramaprasad, dhe Sadler, reagimi është më shumë sesa thjesht informacion. Është informacioni i krijuar në një sistem të veçantë, për një qëllim të veçantë. Kjo është arsyeja se reagimi duhet të jenë specifik për fushën. Për të ndihmuar lojtarët në përmirësim të përqindjes për gjuajtje të lirë, trajnerët e basketbollit jo vetëm u thonë atletëve të gjuajnë topin në rrjetë: ato fokusohen në mekanika të tillë si kujtimi i atletit për tu përkulur në gjuhë dhe për të mbajtur bërërrat të përthyer brenda. Në dhënie e reagimit për nxënësë që janë duke mësuar matematikë, nuk është e dobishme t’i thuhet se duhet të përmirësohen, edhe nëse kjo është e vërtetë. Është më e dobishme të theksohen llojet e gabimeve që janë duke bërë dhe çfarë mund të bëjnë për t’u përmirësuar.

Pika e dytë lidhur me sistemin e reagimit është se është hartuar në mënyrë që informacioni i gjeneruar të jetë i aftë të ndikojë në performancën e ardhme të sistemit. Sapo termi “reagim” është përdorur për të treguar thjesht të dhëna në lidhje me hapësirën ndërmjet performancës aktuale dhe asaj të dëshiruar, ose Akoma më keq, thjesht për përshkrim të nivelit aktual të performancës, humbet çdo lidhje me kuptimin origjinal dhe kuptimplotë.

Kjo është më tepër se një pikë kuptimore. Historikisht, prapashtesa “prapa” ka pasur për qëllim të përshkruajë drejtimin e rrjedhës së informacionit. Në përdorimin e sotëm, duket se përdoret për të treguar kronologji, dhe duket se çdo informacion rreth performancës së mëparshme do të kualifikohet si reagim. Ramaprasad dhe Sadler sugjerojnë se termi “reagim” nuk duhet të përdoret për veç në rast se ka një ndikim mbi performancën në të ardhmen. Të tjerë (si p.sh., Black & Wiliam, 1998b) pranojnë termin “reagim”, por kërkojnë një kusht shtesë, që aktualisht përmirëson të nxënit e nxënësit, për t’u vlerësuar si i pranueshëm. Në çfarëdo mënyre, çfarë është e rëndësishme është pranimi se përdorimi i informacionit të vlerësimit për të përmirësuar të nxënit nuk mund të ndahet nga sistemi mësimor në të cilin ofrohet. Në vitet nga 1986 deri më 1998, nëntë artikuj të rëndësishëm rishikuan ndikimin e praktikave të vlerësimit mbi nxënësit dhe mësimin në kontekstin e klasës, dhe qëndrueshmërinë e gjetjeve nga këto nëntë rishikimet krijuan një interes të veçantë tek kërkuesit, politikë-bërësit dhe praktikantët se si këto ide mund të aplikohen në praktikë. Në kapitullin e mëposhtëm, këto nëntë rishikime, si dhe vepra më bashkëkohore të fushës do të shqyrtohen shkurtimisht.

3. Pasqyrimet e kërkimit mbi vlerësimin dhe mësimin në klasë

Fuchs dhe Fuchs (1986) kanë kryer një meta-analizë të 21 raporteve kërkimore, i cili tregon 96 përmasa të ndryshme të efekteve, në lidhje me nxënës nga klasa parashkollore deri në klasën e dymbëdhjetë, shumica e të cilëve kishin paafësi të moderuara ose të pakta të nxëni. Të gjithë studimet u fokusuan mbi

përdorimin e reagimit për dhe nga mësuesit, me frekuenca ndërmjet dy dhe pesë herë në javë. Përmasa e efektit ishte 0.70 devijime standarde me efekte më të ulëta (0.63) për 22 përmasa e efektit që përfshinin nxënës pa paafësi. Në këto studime ku mësuesit punuan për të vendosur rregulla mbi shqyrtimin e të dhënave dhe veprimeve për t'u ndjekur (rreth gjyca e studimeve të rishikuara),

përmasa e efektit ishte 0.92, ndërkohë kur veprimet u lanë në gjykim të mësuesve, përmasa e efektit ishte vetëm 0.42. Gjithashtu, vlen të përmendet se kur mësuesit morën përsipër hartimin e grafikëve të progresit të fëmijëve si guidë dhe stimul për veprim, efekti ishte më i lartë (0.70) sesa në rastin e kundërt (përmasa e efektit 0.26).

Një tjetër shqyrtim (Natriello, 1987) propozoi një model mbi procesin e vlerësimit i cili konsiston në tetë faza :

1. Përcaktimi i qëllimit të vlerësimit;
2. Përcaktimi i detyrave për nxënësit;
3. Vendosja e kriterëve për performancën e nxënësit;
4. Vendosja e standardeve për performancën e nxënësit;
5. Vëzhgimi i informacioneve mbi performancën e nxënësit;
6. Vlerësimi i performancës së nxënësit;
7. Dhënia e reagimit për nxënësit performues; dhe
8. Monitorimi i rezultateve të vlerësimit të nxënësve.

Konkluzioni i tij kryesor ishte mund të arrihej në përfundim pak nga kërkimi për tre arsye kryesore. I pari ishte se studimet në dispozicion hulumtonin mbi atë çfarë ekzistonte, në krahasim me çfarë mund të ekzistonte, dhe si rezultat kishte prirjen për të konfirmuar problemet me praktikant ekzistuese në vend të dhënies së një baze për praktikën e përmirësuar. I dyti ishte fakti se shumica e studimeve të rishikuara “përqendroheshin në një ose dy aspekte të procesit të vlerësimit. Si rezultat, nuk merrnin në konsideratë ndikimin e elementëve të tjerë kryesorë në përcaktimin e efekteve të vlerësimeve” (f. 170).

Së treti, dhe mbase më i rëndësishëm është fakti se pak nga studimet e rishikuara nga Natriello merrnin në konsideratë në mënyrë të qartë faktin se vlerësimet ishin përdorur në shkolla për një sërë qëllimesh dhe se krahasimet kishin gjasa të ishin çorientuese ku lloje të ndryshme vlerësimesh u krahasuan në terma të funksioneve për të cilat ishin hartuar. Për shembull, një gjetje që diferenconte se reagimi kishte më shumë ndikim mbi të ardhmen e të nxënësit sesa notat që thjesht mund të tregojnë faktin se sistemet në përgjithësi bëjnë në mënyrë më efektive gjërat për të cilat janë hartuar të bëjnë sesa ato që nuk janë hartuar të bëjnë.

Pasqyrimi i tretë (Crooks, 1988) kishte një fokus edhe më të ngushtë – ndikimi i praktikave të vlerësimit në klasa mbi nxënësit. Rishikimi i tij mbulonte vlerësimet formale në klasa si provime, procese informale vlerësimi si pyetje plotësuese në tekste, dhe marrja në pyetje nga mësuesit në klasë. Konkluzioni i tij kryesor ishte se: “Funksionit vlerësues të notave i jepet shumë rëndësi dhe shumë pak rëndësi rolit të tij për të ndihmuar nxënësit në mësim” (f. 468). Ai gjithashtu theksoi se ribalancimi i vëmendjes së kushtuar këtyre dy roleve ishte i nevojshëm, pasi një theksim i tepruar i funksionit të notave jo vetëm që harxhon kohë që mund të përdorej më mirë për aktivitetet e tjera, por në fakt ishte kundërproduktiv, duke rezultuar në:

ulje të motivimit të brendshëm, të ankthit të vlerësimit,tributeve të aftësisë për sukses dhe dështim që dëmton përpjekjet e nxënësve, ul efikasitetin për mësim të nxënësve më të dobët, ulje të përdorimit dhe efikasitetin e reagimit për të përmirësuar mësimin, dhe marrëdhënie shoqërore më të dobëta ndërmjet nxënësve (f. 468)

Studimi i katërt mbi ndikimin e praktikave të vlerësimit në klasë mbi nxënësit dhe mësimi në klasa është bërë nga Bangert-Droëns, Kulik, and Kulik (1991), të cilët

shqyrtuan efektet e testimit të shpeshtë në klasë. Ata zbuluan se nxënësit të cilët kishin të paktën një provim gjatë një periudhe 15-javore rezultuan 0.5 devijime standarde më të larta, dhe se testimet më të shpeshta shoqëroheshin me nivele më të larta arritjeje, megjithëse testimet më shpesh se një herë në dy javë duket se nuk kishte përfitime shtesë. Në një studim të ngjashëm, Bangert-Drowns, Kulik, Kulik, dhe Morgan (1991) raportuan rezultatet e meta-analizës së 58 përmasave të efekteve nga 40 raporte kërkimore mbi efektet e reagimit në evente të quajtura “si testim” si pyetje të vendosura në materialet e programuara mësimore, rishikim i testeve në fund të mësimit, dhe kështu me radhë. Ata arritën në gjetjen se variabla kryesore e ndërmjetme në përcaktimin e ndikimit të reagimit mbi mësimin ishte shkalla në të cilën natyra e reagimit, dhe mënyra se si jepet, inkurajonte “ndërgjegjësimin” tek nxënësit. Kështu, kur nxënësit mund të shikonin përgjigjet përpara dhënies së përgjigjeve, shkalla e arritjes ishte në rënie, por kur studimet kontrolluan mbi këtë “dispozicion para-kërkimit”, përmasa e efektit ishte 0.26 devijime standarde. Për më tepër, kur ndërhyret gjithashtu dhënë reagim për nxënësit në detaje mbi përgjigjen e saktë, në vend të thjesht përgjigjes nëse ishte e saktë apo jo, përmasa e efektit ishte 0.58 devijime standarde.

Këto gjetje u konfirmuan nga Dempster (1991) një rishikim të studimeve mbi efektet e reagimit në provime. Megjithatë ai theksonte se mjaft prej studimeve përkatëse masin arritjen në kuptim të njohurive të lëndës dhe aftësi të nivelit të ulët, kështu që nuk ishte e qartë nëse këto gjetje do të gjeneralizonin në mënyrë të domosdoshme mendimin në shkallë më të lartë. Në një shkrim të mëvonshëm (Dempster, 1992), argumentoi se, ndërsa përfitimet e integritit të vlerësimit në mësim ishin të qarta dhe se ka një konsensus të gjerë në kërkim mbi kushtet për një vlerësim efektiv, duke përfshirë provime të shpeshta pas mësimit, kërkesa akumuluese dhe reagimi pas provimit, vlerësimi neglizhohej në arsimin e mësuesve dhe se kishte evidencë se praktikant aktuale në shkolla ishin larg atyre ideale.

Një rishikim nga Elshout-Mohr (1994), botuar për here të parë në gjuhën holandeze, dhe që pasqyrore shumë studime që nuk janë të disponueshme në gjuhën angleze, sugjeroi se të pasurit njohuri mbi përgjigjet e sakta ishte më e dobishme për detyra të thjeshta por më pak e dobishme për detyra të komplikuar. Korrigjimi i gabimeve mund të jetë një qasje efektive për të nxënësit e materialeve të thjeshta, por për materiale më të komplikuar, të nxënësit kërkon zhvillimin e aftësive të reja, që sipas Elshout-Mohr, kërkon një reagim të një tipi më dialogues, sesa dhënia e përgjigjeve të sakta, dhe për këtë arsye gjithashtu kërkon që nxënësi të jetë aktiv në menaxhimin e procesit.

Shumica e punimit është fokusuar mbi efektet që reagimi ka që shkolla. Megjithatë, në vitin 1996, Kluger dhe DeNisi publikuan një rishikim të efekteve të reagimit në shkolla, kolegje dhe vendet e punës. Ata filluan me përkufizimin e ndërhyrjeve reagimit si “veprime të ndërmarrja nga (një) agjent(ë) i (të) jashtëm për të dhënë informacion lidhur me disa aspekte të performancës së një detyre” (f. 255). Ata identifikuan më tepër se 3000 studime relevante të publikuara ndërmjet viteve 1905 dhe 1995, por vunë në dukje se shumë prej tyre ishin studime të ngushta (në shumë raste kishin vetëm një pjesëmarrës), dhe raportonin detaje jo të mjaftueshme për të lejuar llogaritjen e përmasës së efektit për ndërhyrjen. Për të qenë të sigurt se studimet e një cilësie të dobët nuk ishin të përfshira, Kluger dhe DeNisi përcaktuan tre kriteret për përfshirjen në rishikimin e tyre:

1. Pjesëmarrësit duhet të ndaheshin në dy grupe, dallimi i vetëm ndërmjet grupeve, për aq sa mund të gjykohej, ishte nëse kishin marrë reagim ose jo.
2. Studimi përfshinte të paktën dhjetë pjesëmarrës.
3. Ata përfshinin një llogaritje të performancës me detaje të mjaftueshme për masën e ndikimit që reagimi kishte mbi performancën që do të llogaritej.

Tabela 1

Përgjigjet e mundshme në ndërhyrjet e reagimit (Kluger & DeNisi, 1996).

Lloji i përgjigjes	Reagimi tregon se performanca e tejkalon qëllimin	Reagim tregon se performanca nuk përmbush qëllimin
Ndryshim i sjelljes	Më pak përpjekje	Rritje e përpjekjes
Ndryshimi i qëllimit	Rritje aspirative	Ulje e aspirative
Braktisja e qëllimit	Qëllimi është shumë i thjeshtë	Vendos se qëllimi është shumë i vështirë
Refuzim i reagimit	Reagimi injorohet	Reagimi injorohet

Çuditërisht, vetëm 131 nga 3000 studimet përkatëse përmbushin këto kritere, dhe studimet e përzgjedhura raportuan 607 përmasa të efekteve, duke përfshirë 23,663 vëzhgime të 12,652 pjesëmarrësve. Në të gjitha studimet, përmasa mesatare e efektit për reagim ishte 0.41 devijime standarde, por efektet varonin në mënyrë të konsiderueshme në studime të ndryshme. Veçanërisht, në 50 nga 131 studime (p.sh., 38%), reagimi aktualisht ulte performancën mesatare. Në kërkim të kuptimit të kësaj, ata kërkuar "moderatorët" të efekteve të reagimit dhe zbuluan se ndërhyrjet për reagim ishin më pak efektive kur përqendronin vëmendjen tek vetja, më efektive kur përqendronin vëmendjen në detyrën kryesore dhe më shumë efektive kur përqendroheshin në detajet e detyrës kryesore dhe kur përfshinin përcaktimin e qëllimeve.

Megjithatë, ata arritën në përfundimin se nëse reagimi "funksionon" ose jo, dhe nëse po, sa, nuk ishin pyetjet e duhura për t'i dhënë përgjigje. Ata vunë në dukje se ekzistojnë dy lloje ndërhyrjesh reagimi: ato që tregojnë performancën aktuale që nuk i përket qëllimit të dëshiruar ose performanca aktuale tejkalon qëllimin aktual. Pasi merret, ekzistojnë katër lloj përgjigjesh që individit mund të bëjë – të ndryshojë sjelljen për të arritur qëllimin, të ndryshojë qëllimin, të braktisë qëllimin ose të refuzojë reagimin. Kjo sjell tetë efekte të mundshme të ndërhyrjeve të reagimit, sikurse paraqitet në Tabelën 1.

Me fjalë të tjera, ekzistojnë tetë përgjigje të mundshme për ndërhyrjet e reagimit, dhe gjashtë prej tyre ka qasa të jenë jo efektive ose edhe më keq. Vetëm dy përgjigje, të tërhequra me ngjyrë të zeze në Tabelën 1, kanë gjasa të kenë rezultate pozitive. Por edhe kur shihet se reagimi mund të ketë efekt, mund të mos jetë i qëndrueshëm sikurse është zbuluar nga zbatimet e mësimin me kompjuter. Në rast se reagimi funksionon për të rritur motivimin, atëherë duhet të bëhen përpjekje më të mëdha për të ruajtur motivimin. Reagimi i fokusuar në të nxëniet e detyrave mund të theksojë qëllimet instrumentale, dhe pengon mësimin e thelluar. Në situata të tilla do ishte më mirë të jepej më tepër informacion mbi detyrën madje edhe inkurajimi i një strategjie për përmirësim dhe gjykim, duke dhënë reagim pa një ndërhyrje reagimi. Ata paraqitën se ndërhyrjet e reagimit ishin më pak efektive kur përqendrohen përtej detyrës (për shembull tek vetja) dhe më efektive kur reagimi përqendrohet në motivimin apo të nxëniet e detyrës (kjo do të trajtohet në detaje në diskutimin e rishikimit nga Hattie dhe Timperley dhe në teorinë e "përpunimit të dyfishtë" të Monique Boekaerts si më poshtë). Ky model paraqet gjetjet e mëparshme të njohura që lavdërojnë (sikurse edhe ato që tërheqin vëmendjen nga detyra dhe drejt vetes) kanë shpesh efekte negative (Brophy, 1981).

Black dhe Wiliam (1998a) kërkuar përditësimin e vlerësimeve të diskutuara më lart. Një problem i raportuar ishte vështirësia e përgjithshme në përcaktimin e fushës. Ata vërejtën se shqyrtimet e Natriello dhe Crooks të përmendura më sipër kishin cituar përkatësisht 91 dhe 241 referenca, dhe vetëm nëntë referenca ishin të përbashkëta për të dy shkrimet dhe asnjë nuk kishte cituar shqyrtimin e Fuchs dhe Fuchs. Në punimin e tyre, Black dhe Wiliam paraqitën se përpjekja për tu mbështetur në metodat elektronike ose gjeneroi shumë burime të parëndësishme ose dështoi të identifikonte dokumentet kryesore. Për të qenë të sigurt në shqyrtimin e plotë të fushës, ata shqyrtuan fizikisht çdo numër të 76 revistave që konsiderohen se përmbajnë kërkimet përkatëse të publikuara në vitet 1987 deri 1997. Shqyrtimi i Black dhe Wiliam, që citon 250 studime, arriti në përfundimin se përdorimi i efektiv i vlerësimit në klasë sjell përmirësime në arritjen e nxënësit nga 0.4 dhe 0.7 devijime standarde

(megjithëse vunë re probleme me interpretimin e masave të efektit në studime të ndryshme – shih diskutimin në kapitullin tjetër).

Në shqyrtimin e tyre, Black dhe Wiliam (1998a) fillimisht paraqitën një numër "shembujsh me prova" që ilustruan se numri i veçorive të vlerësimit efektiv formues. Ndoshta tipari më i rëndësishëm ishte se, për të qenë efektiv, vlerësimi formues duhet të integrohej në praktikën e klasës, duke kërkuar një riorganizim thelbësor të veprimeve në klasë:

Është e vështirë të shihet se si çdo risi në vlerësimin formues mund të trajtohet si një ndryshim marginal në punën në klasë. Punë të tilla përfshijnë reagim ndërmjet atyre që mësojnë dhe mësuesit, dhe kjo nënkupton cilësinë e bashkëveprimeve të tyre që është thelbi i pedagogjisë. (Black & Wiliam,

1998a, f. 16) Black dhe Wiliam (1998a) gjithashtu theksuan se që vlerësimi të funksionojë në mënyrë formuese informacion duhet të përdoret dhe kështu trajtimet diferenciale që janë përfshirë në përgjigje të reagimit janë në qendër të mësimin efektiv. Për më tepër, që trajtimet e diferenciale të përzgjidhen në mënyrën e duhur, mësuesve kanë nevojë për modele të përshtatshme rreth mënyrës sesi do të reagojnë nxënësit dhe si të përdorin reagimin. Sikurse Perrenoud (1998) theksoi në komentarin e tij rreth shkrimit të Black dhe Wiliam, "...reagimi i dhënë nxënësve në klasa është si shishet e hedhura në det. Askush nuk mund të jetë i sigurt se mesazhi do të arrijë marrësin" (f. 87). Pasoja e kësaj është se hartimi i vlerësimit formues efektiv nuk mund të shkëputet nga mjedisi mësimor në të cilin është ndërmarrë. Motivimet dhe vetë-perceptimet e nxënësve, dhe historitë e vlerësimit të tyre, do të jenë ndikime të rëndësishme mbi mënyrën sesi reagimi merret (Deci & Ryan, 1994).

Për të adresuar ndikimet në mënyrën e marrjes së reagimit, punimi i Black and Wiliam (1998a) shqyrtoi perspektivën e nxënësit, rolin e mësuesve dhe disa prej sistemeve për organizimin e mësimdhënies në të cilat vlerësimi formues ishte përbërës kryesor. Në hartimin e implikimeve për politikën dhe praktikën e vlerësimit formues, ata arritën në konkluzionin:

Nga ky rishikim nuk del ndonjë model optimal mbi të cilin... mund të bazohet politika. Ajo që shfaqet është një tërësi parimesh udhëzuese, me paralajmërimin e përgjithshëm se ndryshimet në praktikën në klasë që janë të nevojshme janë thelbësore dhe jo marxhinale, dhe duhet të përfshihen nga secili mësues në praktikën e tij/saj sipas mënyrës së tij ose të saj ... Kjo do të thotë se reforma në këtë dimension do të zgjasë në mënyrë të pashmangshme për një kohë të gjatë dhe do të ketë nevojë për mbështetje të vazhdueshme si nga praktikantët, ashtu edhe nga kërkuesit. (f. 62)

Shumica e punës së shqyrtuar më lart u përqendrua në nxënës deri në moshën 18 vjeç. Nyquist (2003) u përqendrua në studimet e reagimit në nxënësit në moshë kolegj. Ai shqyrtoi përafërsisht 3000 studime mbi efektet e reagimit, nga të cilat 86 përmbushin kriteret përfshirëse si më poshtë:

(a) manipulim eksperimental i një karakteristike të rëndësishme për reagimin;

- (b) përdori një grup të nxënësve në moshë kolegjii;
 (c) vlerësoi performancën akademike; dhe
 (d) dha informacion të mjaftueshëm sasior për të llogaritur përmasën e efektit.

Nga 86 studimet ishte e mundur të merren 185 përmasa efekti. Analiza nxori një masë efekti prej 0.40 devijime standard – pothuajse identike me ato të gjetura nga Kluger dhe DeNisi (1996). Peshimi i efekteve në mënyrë që kontributi i tyre në efektin mesatar të ishte proporcional me besueshmërinë e tyre për të zvogëluar këtë efekt mesatar në 0.35 (SE = 0.17), megjithëse vet efektet ishin shumë të ndryshueshme, nga 0.6 deri në 1.6 devijime standarde. Për të hetuar moderatorët e efektit, Nyquist hartoi tipologjinë e mëposhtme të llojeve të ndryshme të vlerësimeve formuese:

Vetëm reagimi i dobët: nxënësve u jepet vetëm njohuri për notën e tyre, shpesh e përshkruar si “njohuri për rezultatet.”

Vetëm reagim: nxënësve u jepet nota së bashku me synime të qarta për të punuar mbi to, ose reagim mbi përgjigjet e sakta për pyetjet, shpesh e përshkruar si “njohuri për rezultatet e sakta”.

Vlerësim formues i dobët: nxënësve u jepet informacion rreth rezultateve të sakta, dhe shpjegim.

Vlerësim formues i moderuar: nxënësit marrin informacion rreth rezultateve të sakta, disa shpjegime, dhe sugjerime specifike për përmirësim.

Vlerësim formues i zhvilluar: nxënësit marrin informacion rreth rezultateve të sakta, dhe aktivitete specifike me qëllim përmirësimin e tyre.

Tabela 2 jep një mesatare të standardizuar të përmasës së efektit për çdo ndërhyrje, megjithatë këto janë vlera të korrigjuara që ndryshojnë në temën origjinale (J.B. Nyquist, komunikim personal, 7 Maj 2007). Rezultatet e Nyquist i bëjnë jehonë gjetjeve të Bangert-Drowns et al. (1991a, 1991b) të diskutuar më lart. Thjesht duke i dhënë reagim nxënësve rreth arritjeve aktuale prodhon vetëm përfitime modeste, por kur reagimi arrin të përfshijë nxënësit në aktivitete të ndërjegjshme, efektet e të nxënësve mund të jenë të mëdha. Përmasat e efektit zbuluar nga Nyquist theksojnë gjithashtu dhe specifikhën e fushës të reagimit efektiv të përmendur më lart.

Nga shqyrtimet e kërkimit të kryer nga Natriello (1987), Crooks (1988), Bangert-Drowns et al. (1991a), Bangert-Drowns et al. (1991b), dhe Black dhe Wiliam (1998a), është e qartë se jo çdo lloj reagimi për nxënësit rreth punës së tyre është njëlloj efektiv. Kërkime bashkëkohore mbështesin këtë pohim. Për shembull, Meisels, Atkins-Burnett, Xue, Bickel, dhe Son (2003) hulumtuan ndikimin e Sistemit të Punës Provë (SPP)—një sistem i vlerësimeve të performancës së përfshirë në kurrikul – mbi arritjet e 96 nxënësve urbanë të klasës së tretë në lexim dhe matematikë, të matura nga Testi i Aftësive Bazë Iowa. Në krahasim me grupin e 116 nxënësve të klasës së tretë në shkollat përkatëse dhe me nxënës në pjesën e mbetur të shkollës së qytetit (Pittsburgh, PA), arritja e nxënësve SPP ishte dukshëm dhe thelbësisht më e lartë në lexim. Megjithatë, në matematikë, diferencat ishin më të vogla dhe nuk arritën në dallime statistikore.

Tabela 2

Përmasa e efektit për lloje të ndryshme të ndërhyrjeve me reagim (Nyquist, 2003).

	N	Përmasa e
efektit		
Vetëm reagim i dobët	31	0.14
Vetëm reagim	48	0.36
Vlerësim formues i dobët	49	0.26
Vlerësim formues mesatar	41	0.39
Vlerësim formues i zhvilluar	16	0.56
Total	185	

Për këtë arsye, lëndë të ndryshme shkollore mund të kërkojnë qasje të ndryshme, duke përfunduar specifikhën e fushës të ndërhyrjeve efektive.

Sistemet e vlerësimit në klasa si Sistemi i Provës së Punës shpesh janë hartuar kryesisht për qëllim përmbledhëse, për të monitoruar dhe raportuar mbi përparimin e nxënësve, për t'u përdorur në prodhim të informacionit për qëllime formuese shpesh të drejtuara për një rol vartësi. Në rishikimin e literaturës kërkimore mbi vlerësimin në klasa, Brookhart (2004) filloi duke ndërmarrë kërkime online me term kërkimi “vlerësimi në klasa”. Duke përfshirë gjetjet që nuk kanë lidhje me edukimin K-12 (për shembull, studime të kryera në mjediset e arsimit të lartë) gjeneroi një total prej 41 studimesh empirike me fokus në punën akademike në arsimin K-12. Ajo arriti në përfundimin se vlerësimi në klasa ndodh në kryqëzimin e tre funksioneve të mësim: mësim, menaxhim i klasës, dhe vlerësim, dhe vuri në dukje se teoria përkatëse për vlerësimin në klasë ishte nga disa fusha të ndryshme, duke përfshirë psikologjinë e dallimeve individuale, studimin e grupeve dhe matjen arsimore. Ajo gjithashtu vuri në dukje se shumë nga studimet që ajo përmendi iu afruan fenomenit në studim nga një perspektivë e vetme disiplinore (shpesh psikologji) ose ishin inventarë jo teorikë të praktikës në klasë. Kur studimet kishin dy ose më shumë këndvështrime praktike ose teorike, ajo arriti në përfundimin se “piktura e vlerësimit në klasë ishte më e pasur dhe shumëdimensionale” (f. 454).

Ndërsa shumë nga studimet e përfshira në rishikimet e diskutuara më lart përqendrohen te nxënësit më të vjetër, është e qartë se qëndrimet e nxënësve për të mësuar formohen nga reagimi që marrin në një moshë shumë të re. Në një studim një-vjeçar të tetë kopshteve dhe klasave të parë në gjashtë shkolla në Angli, Tunstall dhe Gipps (1996a, 1996b) identifikoi një sërë rolesh që luan reagimi. Ashtu si Torrance and Pryor (1998), rezultoi se shumica e reagimit të dhënë nga mësuesi për nxënësit ishte përqendruar në shoqërim: “Unë ndihmoj vetëm ata që kanë ngritur duart lart” (Tunstall & Gipps, 1996b, f. 395). Përtej rolit të shoqërimit, ata identifikuan katër lloje reagimi mbi punën akademike (shiko Tabelën 3). Tipi A përfshin reagim që shpërblen ose ndëshkon nxënësit për punën e tyre, si p.sh. nxënësve u lejohet të largoheshin për drekë herët kur kishin dalë mirë, ose kërcënoheshin se nuk do të lejoheshin për drekë nëse nuk kishin përmbushur detyrat e caktuara. Reagimi i tipit B ishte gjithashtu vlerësues, por, ndërsa reagimi i tipit A përqendrohej në shpërblime dhe dënime, reagimi i tipit B tregonte nivelin e miratimit të mësuesit, p.sh., “Jam shumë i kënaqur” ose “Jam shumë i zhgënjyer nga ty sot”.

Në kontrast me reagimin vlerësues të klasifikuar si tip A dhe B, reagimi i klasifikuar si tip C dhe D është përshkruar. Tipi C fokusohet mbi përshtatshmërinë e punës për sa i përket kriterëve të mësuesve për sukses, duke filluar nga shkalla në të cilën puna tashmë plotëson kriteret (p.sh., “Kjo është plotësisht e shpjeguar”) deri në hapat që nxënësi duhet të ndërmarrë për t'u përmirësuar (p.sh., “Unë dua të shikoni këto dhe të shkruani shenjat tuaja të barazimit në secilën”). Një karakteristikë përcaktuese e reagimit të tipit C të është se përqendrohet në idenë e punës si produkt, ndërsa reagimi i tipit D thekson aspektet e procesit të punës, me mësuesin që luan më tepër rolin e lehtësuesit, sesa të vlerësuesit. Siç shpjegojnë Tunstall and Gipps (1996b) mësuesit e angazhuar në këtë lloj reagimi “përçuan një sens të punës në progres, duke rritur ndërjegjësimin për atë që po bëhej dhe duke reflektuar mbi të” (f. 399).

Nga viti 2002 deri në vitin 2004, si pjesë e programit kërkimor “Çfarë funksionon në risinë në arsim” Organizata për Bashkëpunim dhe Zhvillim Ekonomik (OECD) zhvilloi një rishikim të praktikës së vlerësimit formues në klasat e shkollave të mesme në tetë vende: Australi, Kanada, Danimarkë,

Tabela 3
Tipologjia e reagimit të
mësuesve

Reagim vlerësues	Tipi A	Tipi B
Pozitiv	Shpërbylyes	Miratim
Negativ	Ndëshkues	Refuzim
Reagim përshkrues	Tipi C	Tipi D
	Përcaktim i	Formimi i
Reagim për arritjen	arritjes	arritjes
	Përcaktim i	Formimi i
Reagim për përmirësim	përmirësim	rrugës më tej

Burimi: Adoptuar nga Tunstall and Gipps (1996a).

Angli, Finlandë, Itali, Zelandë e Re dhe Skoci (Looney, 2005). Ashtu sikurse edhe studimet e hollësishtme të rasteve të tetë sistemeve të përfshira në rishikim, raporti i projektit përmbante shqyrtime të kërkimit për vlerësimin formues të botuar në gjuhën franceze (Allal & Lopez, 2005) dhe gjermane (Koller, 2005). Allal dhe Lopez raportuan se puna e studiuesve në Francë dhe vendet frëngjisht-folëse të Belgjikës, Kanadasë dhe Zvicrës janë fokusuar më tepër në punime teorike në krahasim me empirike, me shumë pak studime empirike të kontrolluara. Ato sugjerojnë se gjetja më e rëndësishme e rishikimit të mbi 100 studimeve të botuara në gjuhën franceze gjatë tridhjetë viteve të fundit është se studimet e praktikave të vlerësimit në klasat frëngjisht-folëse kanë përdorur një “koncept të zgjeruar të vlerësimit formues” (f. 245), përgjatë linjave të miratuara nga Black dhe Wiliam (1998a).

Në veçanti, Allal dhe Lopez argumentojnë se koncepti kryesor në qasjen për reagimin, të përfshirë në traditën anglofone, për shembull nga Bloom, është ajo e “përmirësim”, e cila përmbledhet si “reagim + korrigjim.” Në të kundërt, në pjesën më të madhe të kërkimit të ndërmarrë në vende frankofone, koncepti kryesor është “rregullimi”, i përmbledhur si “reagim+përshtatje” (f. 245). Është e rëndësishme të theksohet se fjala frënge *régulation* ka një kuptim shumë më specifik sesa fjala në anglisht “rregullore”. Ka dy mënyra për të përkthyer fjalën “rregullore” në frëngjisht - rëglement dhe rregullore. E para përdoret në kuptim të “rregullave dhe rregulloreve,” ndërsa e dyta përdoret në kuptimin e përshtatjes si në shembullin e termostatit që rregullon temperaturën e dhomës.

Në rishikimin e tyre, Allal dhe Lopez (2005) identifikuan katër zhvillime kryesore në zhvillimin e konceptit të vlerësimit formues në letërsinë në gjuhën frënge gjatë tridhjetë viteve të fundit. Në të parën, të cilin e quajnë “Fokus mbi instrumentalizimin”, theksi ishte në zhvillimin e mjeteve të vlerësimit siç janë bankat e sendeve diagnostikuese dhe sistemet e testimit përshtatës. Në të dytën, quajtur “Kërkimi për korniza teorike”, theksi zhvendoset në një “kërkim për teori që mund të ofrojnë orientim konceptual për kryerjen e vlerësimit” (f. 249). Zhvillimi i tretë—“Studime të praktikave ekzistuese të vlerësimit në kontekstet e tyre” – paraqet bazën për kërkim të kornizave teorike duke e artikuluar atë me studimin se si praktikohet vlerësimi formues në klasat e vërteta. Zhvillimi i katërt dhe më i fundit është “Zhvillimi i pjesëmarrjes aktive të nxënësit në vlerësim” i cili ka shqyrtuar vetëvlerësimin e nxënësit, vlerësimin e bashkëmoshatarëve dhe ndërtimin e përbashkët të vlerësimit nga nxënësit dhe mësuesit.

Nocioni i vlerësimit formues si thelbësor për rregullimin e proceseve mësimore ka marrë miratimin e disa shkrytarëve në komunitetin anglofon (shih, për shembull, Wiliam, 2007), dhe zgjerimi i konceptit të vlerësimit formues në literaturën në gjuhën angleze është vënë në dukje nga Brookhart (2007). Rishikimi i saj i literaturës mbi “vlerësimin formues në klasa” shënoi zhvillimin e konceptit të vlerësimit formues si një serë formulimesh:

Vlerësimi formues jep informacione rreth procesit mësimor;

Vlerësimi formues jep informacione rreth procesit mësimor që mësuesit mund të përdorin për vendime mësimore dhe nxënësit në përmirësimin e performancës së tyre;

Vlerësimi formues jep informacion mbi procesin e mësimdhënies që mësuesit mund të përdorin për vendime të bazuara në udhëzime dhe nxënësit për të përmirësuar performancën e tyre, që motivon nxënësit.

Megjithatë, në përgjithësi, duket se ka më pak lidhje midis punës së fuqishme teorike në traditën frankofone dhe punës empirike të ndërmarrë, veçanërisht në Shtetet e Bashkuara. Allal dhe Lopez (2005) arritën në përfundimin se “studimet e praktikës janë episodike dhe të shpërndara në mjedise të ndryshme, gjë që e bën të vështirë identifikimin e modeleve ose tendencave. Në përfundim, premtimi teorik i punës në gjuhën frënge për vlerësimin formues ka nevojë për një argumentim shumë më empirik” (f. 256).

Rishikimi i literaturës në gjuhën gjermane nga Koller (2005) filloi me një qasje të ngjashme me atë të miratuar nga Black dhe Wiliam, me kërkim të bazave të të dhënave online në të gjithë botimet nga vitin 1980 deri më 2003 nga gjashtë revistat më të rëndësishme në gjuhën gjermane. Koller vuri në dukje se, ndërsa kishte shumë zhvillime në lidhje me vlerësimin formues të raportuar në revista akademike, kishte pak vlerësim të rezultateve të praktikave të vlerësimit formues për nxënësit, megjithëse kishte konfirmime të rëndësishme të disa gjetjeve në literaturën anglofone. Ndoshta më e rëndësishmja, Koller raporton punën e Meyer i cili, ashtu si Kluger dhe DeNisi, zbuloi se ndonjëherë lavdërimi mund të ketë ndikim negativ në të nxënët, ndërsa kritika apo edhe fajësimi mund të jenë ndonjëherë të dobishme. Një pjesë tjetër e rëndësishme e punës së përmendur nga Koller ka të bëjë me ndryshimet ndërmjet mësuesve në përdorimin e normave të referencës. Një numër studimesh, veçanërisht ato të Rheinberg (1980), kanë treguar që nxënësit mësojnë më shumë kur mësohen nga mësuesit që gjykojnë performancën e një nxënësi kundrejt performancës së mëparshme të të njëjtit nxënës (një normë referimi individuale) sesa mësuesit që krahasojnë nxënësit me të tjerët në klasë (një normë referimi sociale)

Kohët e fundit, janë shfaqur tre rishikime thelbësore për vlerësimin formues. E para (Wiliam, 2007), përqendrohet posaçërisht në mësimin e matematikës.

Ashtu sikurse shqyrtimi i provave të kërkimit mbi vlerësimin formues, Wiliam nxori disa nga implikimet e kërkimit për mësimin e matematikës dhe përshkroi se si idetë qendrore të vlerësimit formues mund të integrohen brenda idesë më të gjerë të rregullimit të të mësuarit procese të zhvilluara nga përmbledhja e letërsisë në gjuhën frënge të përmendura më lart.

Dy rishikimet më të fundit u shfaqën në vitet e ardhshme në revistën Rishikimi i Kërkimit Arsimor. Si pjesë e një programi kërkimor më të gjerë mbi zhvillimin e mjediseve arsimore inteligjente, Shute (2008) shqyrtoi kërkimin mbi reagimet për nxënësë. Gjithsej 141 botime plotësuan kriteret e përfshirjes (103 artikuj revistash, 24 libra dhe kapituj librash, 10 konferenca dhe katër raporte kërkimore). Ndërsa, siç pritej, shqyrtimi i Shute identifikoi boshllëqe të mëdha në literaturë dhe arriti në përfundimin se nuk ekziston një përgjigje e thjeshtë për pyetjen. “Çfarë lloj reagimi funksionon?”, rishikimi miratoi gjetjet e pasqyrimeve të mëparshme mbi masën e efekteve që mund të priten nga reagimet: efektet e standardizuara të masave nga 0.4 në 0.8 devijime standard. Shute gjithashtu ofroi një numër udhëzimesh paraprake për hartimin e një reagimi efektiv:

1. Udhëzime për të përmirësuar të nxënësit. Reagimi duhet të përqendrohet në veçoritë specifike të detyrës, dhe të ofrojë sugjerime se si të përmirësohet, në vend të përqendrimit të nxënësit; duhet të përqendrohet në “çfarë, si dhe përse” të një problemi sesa thjesht t’u tregojë nxënësve nëse ishin të saktë apo jo; reagim i shtjelluar duhet të paraqitet në njësi të menaxheshme dhe, duke i bërë jehonë fjalës së famshme të Ajnshtajnit duhet të jetë “sa më i thjeshtë të jetë e mundur, por jo më i thjeshtë.”. Megjithatë, reagimi nuk duhet të jetë aq i hollësishëm dhe specifik sa të ekspozojë mësimin plotësisht aq sa nxënësit nuk kanë nevojë të mendojnë vetë. Reagimi është gjithashtu më i efektshëm kur është nga një burim i besueshëm (qoftë njerëzor apo kompjuter).

2. Udhëzime në lidhje me kohën e reagimit. Koha optimale e reagimit duket se varet fuqimisht nga lloji i mësimin. Një reagim i menjëhershëm duket se është më i dobishëm për mësimin procedural, ose kur detyra është përtëjtë aftësive të nxënësit në fillim të mësimin, ndërsa reagimi i vonuar duket se është më i përshtatshëm për detyrat brenda aftësisë së nxënësit, ose kur transferimi të tjetri kërkon kontekst.

Një rishikim i Hattie dhe Timperley (2007) përmbledh një program të gjerë të punës së realizuar nga Hattie dhe kolegët e tij mbi vlerësimin sistematike të ndikimeve në arritjet e nxënësve. Një punim i mëparshëm (Hattie, 1999) përshkroi ndërtimin e një baze të dhënash prej 500 meta-analizash, duke raportuar 450,000 përmasa të efektit nga 180,000 studime që përfshijnë mbi 20 milion pjesëmarrës. Një analizë e 74 meta-analizash përdorur në studimin e vitit 1999 që përmend në mënyrë specifike reagimin gjeti një efekt mesatar të masës së lejuar 0.56 në 13,370 përmasa të efektit në 74 meta-analiza (Hattie & Timperley, 2007), por Hattie dhe Timperley arritën në përfundimin, sikurse Kluger dhe DeNisi (1996), se ka ndryshueshmëri të konsiderueshme ndërmjet studimeve të reagimit të efektet e tyre në të nxënë. Mesatarja e 5755 studimeve mbi përmasat e efektive që Hattie dhe Timperley përmbledhën si “reagim” ishte 0.95 devijime standarde, të mbështetura nga vetëm 89 studime të kodifikuara si “orientime”, që mesatarisht është 1.1 standard devijimi.

Hattie dhe Timperley përcaktojnë qëllimin e reagimit si zvogëlim i mospërputhjeve ndërmjet kuptimeve aktuale ose performancës dhe një qëllimi të dëshiruar (të propozuar nga Ramaprasad, 1983). Duke u bazuar në punën e Deci dhe Ryan (1994) dhe Kluger dhe DeNisi (1996), modeli i tyre sugjeron që nxënësit mund të ulin mospërputhjet qoftë duke zbatuar strategji më efektive ose duke rritur përpjekjet ose duke braktisur, megjulluar ose ulur qëllimet që kanë vendosur për veten e tyre. Mësuesit mund të zvogëlojnë mospërputhjen duke ndryshuar vështirësinë ose specifikën e qëllimeve, ose duke dhënë më shumë mbështetje për nxënësin. Modeli specifikon tre lloj pyetjesh që reagimi është hartuar të japë përgjigje (Ku jam duke shkuar? Si jam duke shkuar? Çfarë vjen më pas? dhe çdo pyetje reagimi funksionon në katër nivele: reagim mbi detyrën (FT), reagim mbi kryerjen e detyrës (FP), reagim mbi vetërregullimin (FR) dhe reagim për veten si person (FS). Ata demonstrojnë se FS është lloji më pak efektiv, pasi FR dhe FP “janë më të fuqishëm në termat e përpunimit më të thellë dhe mjeshhtërisë së detyrave” (f. 90-91) ndërsa FT është i fuqishëm kur reagimi përdoret ose për të përmirësuar përpunimin e strategjisë, ose për të përmirësuar vetërregullimin (megjithatë theksojnë se këto kushte rrallë zbatohen në praktikë). Roli i vetë rregullimit në vlerësimin formues shtjellohet në detaje si më poshtë.

4. Përkufizimet e vlerësimit formues dhe vlerësimit për të nxënësit

Ndërsa kërkimi i shqyrtuar më lart sugjeron se përdorimi i vlerësimit për të informuar udhëzimet mund të ketë ndikim të rëndësishëm në mësim, vlerësime të ndryshme gjejnë përmasa të efektive të ndryshme

përfitimet e vlerësimit formues. Kluger dhe DeNisi (1996) paraqitën një përmase efekti mesatar me 0.41 për ndërhyrje reagimi, ndërsa Black dhe Wiliam (1998a, 1998b) llogaritën se efektet e vlerësimit formues ishin rreth 0.4-0.7 devijime standarde. Shute (2008) sugjeroi një rang të ngjashëm (0.4-0.8), por Hattie dhe Timperley propozuan një përmase efekti mesatar prej 0.96 devijime standard për efektin e reagimit. Nga ana tjetër, në një mjedis klase, të zhvilluar gjatë një viti, me mësues të zakonshëm, dhe ku performanca u mat duke përdorur testime standarde të mandatuara nga jashtë, Wiliam, Lee, Harrison, dhe Black (2004) konstatuan se një sërë strategjish vlerësimi formues të prezantuara nga mësuesit kishte një përmase efekti prej 0.32 devijime standarde. Ky është një efekt thelbësor (autorët vlerësuan se kjo ishte ekuivalente me një rritje të shkallës së mësimin të nxënësve prej 70%, ose tetë muaj shtesë mësimin për një vit), por vetëm një e treta e masës së efektive të sugjeruara nga Hattie dhe Timperley.

Pjesë e ndryshueshmërisë është, pa dyshim, e shkaktuar nga ndryshimet në ndryshueshmëri e masave të përdorura në studime të ndryshme ndaj efektive të mësimin - shih Wiliam (2010 f. 20-22) për një diskutim mbi marrëdhëniet ndërmjet ndryshueshmërisë ndaj udhëzimit dhe përmases së efektit. Përmase e efektit gjithashtu do të preket nga ndryshimet në mospërputhje të popullsisë. Shumë studime të përfshira në rishikimet e kërkimit janë kryer në nën-popullata që nuk përfaqësojnë të gjithë popullsinë. Për shembull, nëse përmase e efektit është llogaritur në një studim të ndërhyrjeve të ndryshme për nxënë me nevoja të veçanta studimi, atëherë ajo përmase efekti nuk do të jetë e përgjithshme për të gjithë popullsinë—kur popullsia është më e ndryshueshme, devijimi standard që përdoret si emërues në llogaritjen e përmases së efektit është më i madh, duke çuar në një vlerësim më të vogël të përmases së efektit.

Megjithatë, duket se një pjesë e rëndësishme – mbase edhe shumica e ndryshueshmërisë shkaktohet nga dallimet në mënyrën se si idetë e vlerësimit formues ose vlerësimit për të mësuar ishin operacionale. Siç vë në dukje Bennett (2009), në një studim kritik të rëndësishëm të fushës, nuk mund të jemi të sigurt për efektet e këtyre ndryshimeve në praktikë përveç nëse dikush ka një përkufizim adekuat të asaj që kuptojnë termat vlerësimi formues dhe vlerësimi për të nxënëit, dhe një lexim i ngushtë i përkufizimeve që janë dhënë sugjeron që nuk ka dakordësi të qartë për kuptimet e termave vlerësim formues dhe vlerësim për të nxënëit.

Sikurse përmendet më sipër, Bloom koncepton vlerësimin formues si një kombinim të reagimit dhe korrigjimeve udhëzuese. Black dhe Wiliam (1998a) përcaktojnë vlerësimin formues si më poshtë:

Përdorim termin e përgjithshëm vlerësim për t’iu referuar të gjithë aktiviteteve të ndërmarra nga mësuesit – dhe nga nxënësit e tyre për të vlerësuar veten e tyre – që japin informacione që do të përdoren si reagimit për të ndryshuar mësimdhënien dhe aktivitetet mësimore. Një vlerësim i tillë behët vlerësim formues kur provat përdoren për të përshtatur mësimin në përmeshje të nevojave të nxënësve” (Black & Wiliam, 1998b, f. 140)

Një numër autorësh kanë propozuar përkufizime disi më të ngushta, më së shpeshti duke kërkuar që ndryshimet në mësim të ndodhin gjatë mësimin, sikurse ilustron katër citatat e mëposhtme:

“procesi i përdorur nga mësuesit dhe nxënësit për të njohur dhe për t’iu përgjigjur të nxënëit e nxënësit me qëllim përmirësimin e të nxënëit, gjatë mësimin” (Cowie & Bell, 1999, f. 32)

“vlerësimi i kryer gjatë procesit mësimor me qëllim përmirësimin e mësimdhënies apo mësimin” (Shepard et al., 2005, f. 275)

“Vlerësimi formues i referohet vlerësimin të shpeshtë, interaktiv përgjigje informacionit me dialog, demonstrim dhe të progresit dhe të kuptuarit të nxënësve për të identifikuar nevojat mësimore dhe rregullimin e mësimdhënies në mënyrën e duhur” (Looney, 2009, f. 264) 2005, f. 21)

“Një vlerësim formues është një mjet që mësuesit përdorin për të matur kuptimin e termave specifikë dhe aftësive që i shpjegojnë nxënësve. Është një mjet ‘i ndërmjetëm’ për të identifikuar konceptet e gabuara të nxënësve dhe gabimet gjatë mësimin të materialit (Kahl, 2005, f. 11)

Grupi i Reformës së Vlerësimin—një grup i përkushtuar për të siguruar që politika dhe praktika e vlerësimin informohen nga prova të kërkimit – pranuan ndikimin që vlerësimi ka në të nxënët, për mirë dhe për keq, dhe propozuan shtatë rregulla që përmbledhin karakteristikat e vlerësimin që promovojnë të nxënët:

është mishëruar në funksion të mësimdhënies dhe të nxënët në të cilën është pjesë thelbësore;
përfshin ndarjen e qëllimeve të të nxënët me nxënësit;
synon të ndihmojë nxënësit të dinë dhe të njohin standardet që synojnë;
përfshin nxënësit në vetëvlerësim;
siguron reagim gjë që i bën nxënësit të njohin hapat e tyre të ardhshëm dhe si t’i ndërmarrin ato;
mbështetet në besimin se çdo nxënës mund të përmirësohet;
përfshin mësuesit dhe nxënësit të rishikojnë dhe reflektojnë mbi të dhënat e vlerësimin (Broadfoot et al., 1999, f. 7).

Në kërkimin e një termi për të përshkruar vlerësime të tilla, ata sugjerojnë që termi vlerësim formues të përdorej në mënyra të ndryshme që nuk ishte më e dobishme:

Vet termi ‘formues’ është i hapur për një sërë interpretimesh dhe shpesh nënkupton jo më shumë se vlerësimi që kryhet shpesh dhe planifikohet në të njëjtën kohë me mësimdhënien. Një vlerësim i tillë nuk ka domosdoshmëritë të gjitha karakteristikat e identifikuar si ndihmë për mësimin.

Mund të jetë formues në ndihmë të mësuesit për të identifikuar fushat ku nevojitet më tepër shpjegim ose praktikë. Por për nxënësit, notat apo komentet mbi punën e tyre mund të tregojnë suksesin ose dështimin, por jo progresin e bërë për mësim të mëtejshëm. (Broadfoot et al., 1999, f. 7)

Në vend të kësaj, ata preferuan termin vlerësim për të nxënët, të cilin e përkufizuan si “procesi i kërkimit dhe interpretimit të provave për përdorim nga nxënësit dhe mësuesit e tyre për të vendosur se ku janë nxënësit në të nxënët e tyre, ku duhet të shkojnë dhe si mund ta arrijnë këtë sa më mirë” (Broadfoot et al., 2002, fq.2-3).

Përdorimi më i hershëm i termit vlerësim për të nxënët është një kapitull i atij titulli të Harry Black (1986). Ishte gjithashtu titulli i shkrimit të dhënë nga AERA në vitin 1992 (James, 1992)—në të njëjtin vit që libri i titulluar Testim për të nxënët u botua në SH.B.A (Mitchell, 1992)—dhe tre vjet më vonë, si titulli i një libri nga Ruth Sutton (1995). Në Shtetet e Bashkuara, origjina e termit shpesh i akordohet gabimisht Rick Stiggins si rezultat i popullarizimit të termit (shih, për shembull, Stiggins, 2005), megjithëse vet Stiggins gjithmonë ia ka atribuar termin autorëve të tjerë.

Kohët e fundit, një konferencë ndërkombëtare për vlerësimin për të nxënët në Dunedin në vitin 2009, duke u bazuar në punën e bërë në dy konferenca të mëparshme në Mbretërinë e Bashkuar (2001) dhe SHBA (2005), miratoi përkufizimin e mëposhtëm:

Vlerësimi për të nxënët është pjesë e praktikës së përditshme nga nxënësë, mësues dhe kolegë që kërkojnë, reflektojnë dhe i japin

observim në mënyra që përmirësojnë mësimin (Klenowski,

Termi vlerësimi për të nxënët ka padyshim interes, veçanërisht kur është në kundërshtim me vlerësimin për të nxënët, por sikurse Bennett (2009) thekson, zëvendësimi i një termi me një tjetër shërben thjesht për të lëvizur barrën përcaktuese. Më e rëndësishmja, siç kanë theksuar Black dhe Wiliam dhe kolegët e tyre, dallimet ndërmjet vlerësimin për të nxënët dhe vlerësimin të mësimin nga njëra anë dhe vlerësimin formues dhe përmbledhës nga ana tjetër janë të ndryshme në natyrë. Dallimi lidhet me qëllimin për të cilin bëhet vlerësimi, ndërsa i dyti ka të bëjë me funksionin që në fakt shërben:

Vlerësimi për mësimin është çdo vlerësim për të cilin prioritet në hartimin dhe praktikën e tij është t’i shërbejë qëllimit për të përmirësuar të nxënët e nxënësve. Në këtë mënyrë ndryshon nga vlerësimi i hartuar kryesisht për të shërbyer qëllimeve të përgjegjësisë, ose të renditjes ose të vërtetimit të kompetencës. Një aktivitet vlerësimi mund të ndihmojë të nxënët nëse ofron informacione që mësuesit dhe nxënësit e tyre mund të përdorin si reagim në vlerësimin e vetes dhe të njëri – tjetrit dhe në modifikimin e aktiviteteve mësimore në të cilët janë angazhuar. Një vlerësim i tillë bëhet “vlerësim formues” kur provat aktualisht përdoren për të përshtatur punën mësimore për të përmbushur nevojat e të nxënët. (Black, Harrison, Lee, Marshall, dhe Wiliam, 2004, f. 10)

Bennett (2009) mbështet idenë se barazimi i vlerësimin për të nxënët me vlerësimin formues dhe vlerësimin e mësimin me vlerësimin përmbledhës është i padobishëm dhe i thjeshtëzuar. Për një pikëpamje “më të larmishme” (f. 5) si sugjeroi që vlerësimet e hartuara kryesisht për t’i shërbyer një funksioni përmbledhës mund të funksionojnë në mënyrë formuese, ndërsa ato të hartuara kryesisht për t’i shërbyer një funksioni formues mund të funksionojnë gjithashtu si përmbledhje. Konsideroni shtatë skenarët e vlerësimin në vijim.

1. Një ekip mësuesish të së njëjtës shkollë takohen për të diskutuar nevojat e zhvillimit të tyre profesional. Ata analizojnë nevojat e tyre për zhvillim profesional. Pasi analizojnë notat

e marra nga nxënësit në teste kombëtare vunë re se ndërsa rezultatet e tyre janë, përgjithësisht, të krahasueshme me standardet kombëtare, nxënësit kanë tendencën të kenë rezultate më të ulta në fusha që përfshijnë përpjesëtim dhe proporcion. Ata vendosën në fokus të aktiviteteve të tyre të zhvillimit profesional përpjesëtimin dhe proporcionin për vitin e ardhshëm, duke u takuar rregullisht për të diskutuar ndryshimet që bënë në mënyrën e mësimin të kësaj teme. Dy vjet më vonë, ata zbulojnë se nxënësit e tyre po shënojnë rezultate të mira në përpjesëtim dhe proporcion në provimet kombëtare, gjë që rrit rezultatet e nxënësve mbi standardet kombëtare.

2. Çdo vit, një grup mësuesish të klasës së katërt takohen për të rishikuar performancën e nxënësve mbi një provim leximi standard dhe në veçanti, shikojnë objektin (proporcionin e saktë) për lloje të ndryshme të artikujve në provim. Kur objektet e artikujve janë më të ulëta nga sa pritej, ata shikojnë sesi u planifikua dhe u dha udhëzimi në lidhje me atë aspekt të leximit, dhe shikojnë mënyrat në të cilat udhëzimi mund të përmirësohet vitin e ardhshëm.

3. Çdo shtatë javë, mësuesit e një shkollë përdorin një sërë provimesh “të ndërmjetme” për të kontrolluar përparimin e nxënësve. Çdo nxënës që merr rezultate më të ulëta se kufiri që gjykohet i nevojshëm për të përparuar ftohet për të qenë pjesëmarrës në mësim shtesë. Çdo nxënës që ka rezultate nën kufi në dy raste të njëpasnjëshme ka të nevojshme të marrë pjesë në mësim shtesë.

4. Një mësues harton një njësi udhëzuese mbi levën ngritëse. Duke ndjekur modelin që është i përhapur në shkollat e mesme në Japoni

(Lewis, 2002 f. 76), megjithëse 14 periudha janë ndarë në njësi, mësuesi sigurohet që e gjithë përmbajtja të jetë e përfshirë në 11 periudhat e para. Në periudhën 12, nxënësi jep një provim mbi 11 periudhat e mëparshme, dhe mësuesi mbledh përgjigjet e nxënësve, i lexon ato, dhe mbi bazë të asaj çka mësuar rreth kuptimit të lëndës në klasë, planifikon se çfarë do të bëjë në mësimet 13 dhe 14.

5. Një mësuesi sapo ka diskutuar në klasë përse dokumentet historike nuk mund të merren në vlerën e tyre. Ndërsa mësimi përfundon, çdo nxënësi i jepet një kartë indeksi (8 cm x 13 cm) dhe i kërkohet t'i japë përgjigje pyetje "Përse historianët janë të shqetësuar për paragjykimet në burimet historike?" Ndërsa largohen nga klasa, nxënësit i dorëzojnë mësuesit "kartat e daljes" dhe pasi të gjithë nxënësit janë larguar, mësuesi lexon kartat dhe vendos se si do të nisë mësimin e ardhshëm.

6. Një klasë e gjashtë ka mësuar rreth llojeve të ndryshme të gjuhës figurative. Në mënyrë që të kontrollohet kuptimi i klasës, mësuesi i jep çdo nxënësi një set me gjashtë karta që përmbajnë shkronjat A, B, C, D, dhe E. Në tabelën e bardhë interaktive, tregon listën e mëposhtme:

- A. Aliteracion
- B. Onomatopi
- C. Hiperbolë
- D. Personifikimi
- E. Krahasimi

Më pas lexohen një sërë deklaratash:

1. Ai ishte si një dem në një dyqan kinez.
2. Kjo çantë peshon një ton.
3. Ai ishte i gjatë sa një shtëpi.
4. Buzëqeshja e ëmbël vezulluese. . .
5. Ai i ra burisë për çiklistin.

Ndërsa lexohet çdo deklaratë, secili duhet të mbajë kartat për të treguar se cilën figurë letrare ka dëgjuar. Mësuesi/ja kupton se pothuajse të gjithë nxënësit kanë supozuar se çdo fjali mund të ketë vetëm një figurë letrare. Fjalja e tretë është krahasim, por gjithashtu hiperbolë, dhe më pas dy deklaratat e fundit duhet të rizgjidhen nga klasa. Për më tepër, mësuesi/ja shënon në mendje tre nxënësit që iu janë përgjigjur gabim shumicës së pyetjeve, në mënyrë që të mund të vijojë me ta individualisht në një moment tjetër.

7. Një mësuesi kimie i shkollës së mesme ka shpjeguar një klasë se si të barazojë ekuacionet kimike. Për të testuar klasën, ajo shkruan ekuacionin e pabarazuar për reaksionin e hidroksidit merkurit me acidin fosforik. Më pas ajo fton nxënësit të ndryshojnë sasinë e elementëve të ndryshëm në ekuacion, dhe kur nuk ka më sugjerime nga klasa, ajo pyet klasën për të votuar nëse ekuacioni është i saktë. Të gjithë pohojnë. Mësuesi/ja arrin në përfundimin se klasa e ka kuptuar dhe vazhdon më tej.

Në secilin prej këtyre shtatë skenarëve, informacioni i vlerësimit është përdorur për të marrë një vendim më të mirë në lidhje sesa do të ishin marrë në mungesë të provave. Në dy skenarët e parë, instrumenti i përdorur i vlerësimit ishte hartuar tërësisht në shërbim të një funksioni përmbledhës, por mësuesit e përfshirë gjetën një mënyrë për të përdorur provat rreth arritjeve të nxënësit të nxitur nga vlerësimi për të përmirësuar udhëzimin e tyre.

Në gjashtë skenarët e parë, përdorimi ndryshoi mësimin për më mirë ndërsa gjatë të fundit, informacioni i vlerësimit konfirmoi se ajo që mësuesi kishte planifikuar të bënte ishte veprimi i duhur.

Në këtë kuptim, ishte vendim më i mirë në krahasim me atë

me çfarë do ishte në mungesë të provave pasi ishte më i bazuar. Me fjalë të tjera, prova nga vlerësimet u përdorën për përmirësimin e të nxënësit, megjithëse shumë nga autorët e përmendur më lart, me siguri nuk do t'i konsiderojnë tre të parët si vlerësim për të nxënësit ose vlerësim formues.

Një përgjigje për këtë do të ishte për përpjekja për të kufizuar kuptimin e vlerësimit formues ose vlerësimit për të nxënësit në llojet e vlerësimit që janë afër udhëzimit, që do të përjashtonte tre skenarët e parë dhe për disa autorë gjithashtu të katërtin. Megjithatë, kufizimi i kuptimit apo përdorimit të termave vlerësim për të nxënësit dhe vlerësim formues thjesht do të siguronte se termat zbatohen vetëm në praktika që konsiderohen të favorshme është përkundrazi gabim.

Është si qasja e përdorur nga personazhi i Lewis Carroll në Njëpërmjet xhamit: "Kur unë përdor një fjalë . . . do të thotë atë që unë zgjedh të thotë – as më pak as më shumë" (Carroll, 1871). Vlera e përdorimit të termit vlerësim në fraza si vlerësim për të nxënësit dhe vlerësim formues tërheq vëmendje për faktin se proceset në konsideratë mund të shikohen si procese vlerësimi.

Për këtë arsye, duket më e dobishme të pranojmë se në secilën prej këtyre rasteve, vlerësimi është kryer me qëllim të përmirësimit të të nxënësit (megjithëse kjo mund të mos ketë qenë arsyeja e vetme për vlerësimin), dhe se dëshmitë nga vlerësimi janë përdorur për të përmirësuar mësimin. Për këtë arsye, Black dhe Wiliam përsëritën përkufizimin e tyre origjinal në një mënyrë paksa të ndryshme, për të cilën ata sugjeruan se ishte në përputhje me përkufizimin e tyre origjinal dhe të tjerët të dhënë më lart, përfshirë atë të Grupit të Reformës së Vlerësimit:

Praktika në një klasë është formuese në atë masë që provat në lidhje me arritjet e nxënësve nxiten, interpretohen dhe përdoren nga mësuesit, nxënësit ose kolegët e tyre, për të marrë vendime rreth hapave të mëtejshëm në mësim që ka gjasa të jenë më të mirë, ose të themeluar më mirë, sesa vendimet që do të ishin marrë në mungesë të provave të përzgjedhura. (Black & Wiliam, 2009, f. 9)

Hulumtimi i plotë i pasojës së këtij përkufizimi është përtej qëllimit të këtij punimi, por ky përkufizim kërkon sqarime. Në shpjegimin e këtij përkufizimi, Black dhe Wiliam sqarojnë se termi "udhëzim" është përdorur në kuptimin në të cilin përdoret në Shtetet e Bashkuara – modeli i mjediseve të mësimi—dhe "hapat e ardhshëm në mësim" mund të merren nga mësuesit, nxënësit, ose kolegët e tyre, ose ndonjë kombinim i të treve. Arsyetimi për secilin prej komponentëve të përkufizimit mund të gjendet në Black dhe Wiliam (2009) dhe Wiliam (2010), dhe shqyrtimet se si vlerësimi formues lidhet me perspektivat e tjera teorike mbi mësimdhënien dhe mësimin mund të gjenden në Black and Wiliam (2004, 2011) dhe Wiliam (2007). Në pjesën e fundit të këtij punimi, shqyrtohen disa nga kushtet që duhet të ekzistojnë për vlerësimin me qëllim të nxënësit.

5. Kur vlerësimi mbështet mësimin?

Ndërsa përkufizimi i propozuar nga Black dhe Wiliam është relativisht i saktë, është më tepër një mjet për të përcaktuar nëse një vlerësim ka funksionuar në mënyrë formale sesa është një recetë për të gjeneruar vlerësime që do të jenë ose ka të ngjarë, të funksionojnë në mënyrë formale. Nga kërkimi i studiuar më sipër, shfaqen dy karakteristika të rëndësishme në hartimin e vlerësimit që mbështet studimin. Njëra është prova që gjenerohet është "udhëzim i gjurmueshëm" që mbështet të nxënësit. (Wiliam, 2007). Me fjalë të tjera, prova është më shumë se informacion mbi praninë e boshllëkut

	Ku shkon nxënësi	Ku është nxënësi	Si të arrijë atje
Mësues	Sqarimi i synimeve të të nxënësve dhe shpërndarja dhe kriteret për sukses	Diskutime efektive në klasë, Duke dhënë reagim që shtyn nxënësit të aktivitetit dhe detyra që arrijnë në prova të mësimin	Duke dhënë reagim që shtyn nxënësit përpara
Bashkëmoshatar	Të kuptuarit dhe shpërndarja e synimeve të leximit dhe kriteret për sukses	Aktivizimi i nxënësve si burime udhëzuese për një tjetër	
Nxënës	Kuptimi i synimeve të mësimin dhe kriteri për sukses	Aktivizimi i nxënësve si pronarë të mësimin të tyre	

Fig. 1. Aspektet e vlerësimin formues (Wiliam & Thompson, 2008).

ndërmjet performancës aktuale dhe të dëshiruar. Prova duhet të japë informacione mbi llojet e aktiviteteve udhëzuese ka të ngjarë të rezultojë në përmirësim të performancës. Për shembull, një rezultat i ulët në matematikë ka të ngjarë të tregojë asgjë më shumë se një nxënës nuk ka mësuar ende se çfarë kishte për qëllim të mëson. Udhëzimi i vetëm që i jepet mësuesit është se nevojiten më shumë udhëzime. Nëse vlerësimi është krijuar për të mbështetur konkluzione të vlefshme në lidhje me aspektet specifike të performancës, atëherë mësuesi mund të kuptojë gjithashtu që nxënësi ka vështirësi të veçanta me fraksionet e renditjes së gradave. Kjo bën të mundur që mësuesi të përqendrojë udhëzime përmirësuese, por gjithashtu shpjegon se përse nxënësi ka vështirësi. Megjithatë, nëse vlerësimi zbulon një çështje specifike – për shembull që nxënësi beson se vetëm masa e emëruesit ka rëndësi në krahasim me fraksionet (Vinner, 1997) atëherë kjo siguron udhëzime të qarta për mësuesin në lidhje me llojet e aktiviteteve të vetëvlerësimin për nxënësin.

Kërkesa e dytë ka të bëjë me faktin që nxënësi të ndërmarr veprime për të përmirësuar të nxënësve; kjo mund të përfshijë ndërmarrjen e veprimeve përmirësuese të ofruara nga mësuesi, duke i kërkuar një bashkëmoshatar ndihmë, ose duke reflektuar mbi mënyrat e ndryshme për të ecur përpara me mësimin – në fund të fundit, reagimi i hartuar më mirë është i padobishëm nëse nuk veprohet për të. Me fjalë të tjera, reagimet nuk mund të vlerësohen pa marrë parasysh kontekstin e realizimit në të cilin është siguruar dhe përdorur. Në të njëjtën mënyrë që inxhinierët hartojnë sisteme reagimit sesa thjesht mënyra të thjeshta për të gjeneruar të dhëna, për të kuptuar reagimin duhet të shikojmë mjedisin mësimor në të cilin vepron. Një reagim i mirëmunduar dhënë nxënësve të cilët besojnë se nuk janë “të mirë” në një lëndë të caktuar ka të ngjarë të injorohet ose refuzohet, ose të përshtatet në një mënyrë tjetër për të lejuar nxënësit të ruajë ndenjen e mirëqenies. Përfshirja e nxënësve dhe bashkëmoshatarëve të tyre ishte përfshirë në mënyrë të qartë nga Wiliam and Thompson (2008) në propozimin e tyre se vlerësimi formues mund të konceptohet se përfshin tre procese kryesore (duke identifikuar se ku janë nxënësit në të nxënësve, ku po shkojnë dhe si mund të arrijnë atje) ushtruar nga tre kategori aktorësh (mësues, nxënës, bashkëmoshatar).

Ata sugjerojnë se matrica rezultuese e nëntë qelizave mund të organizohet në pesë “strategji kryesore” të vlerësimin formues, siç tregohet në

Fig. 1.

Ndërsa secila prej këtyre pesë “strategji kryesore” ka krijuar një bazë thelbësore kërkimi (shih Wiliam, 2007, për përmblendje) ato gjithashtu mund të shihen kolektivisht si strategji për rregullimin e proceseve të të nxënësve. Në të vërtetë, Black and Wiliam (2009) theksojnë se vlerësimi formues përfshin “krijimin dhe kapitalizimin mbi “momentet e emergjencës” në udhëzime me qëllim të rregullimit të proceseve mësimore” (f. 12).

Çdo përpjekje për rregullimin e proceseve mësimore kërkon të paktën disa ide qëllimi, qoftë ky është konceptuar si një destinacion i vetëm për të nxënësve, ose një “horizont” i gjerë i qëllimeve mësimore cilido prej të cilave është pranuar në mënyrë të barabartë. Roli i mësuesit është të nxjerrë prova të arritjeve dhe ndërmarr veprimet e duhura për të drejtuar të nxënësve në drejtimin e synuar. Me këtë formulim, roli i bashkëmoshatarëve është i ngjashëm me atë të mësuesve – ndërsa

bashkëmoshatarëve mund të kenë mangësi në trajnim dhe eksperiencë të mësuesve, ata kanë njohuri unike mbi të nxënësve, dhe për shkak të marrëdhënieve ndërmjet bashkëmoshatarëve janë të ndryshme nga ato ndërmjet mësuesve dhe nxënësve, do të kenë strategji të hapura që nuk do të ishin të hapura, ose do të ishin më pak efektive, kur përdoren nga mësuesit.

Strategjia përfundimtare, “Aktivizimi i nxënësve si pronarë të të nxënësve të tyre” tregon në mënyrë të qartë një numër fushash kërkimi, të tilla si metakognitiv (Hacker, Dunlosky, & Graesser, 1998), motivimi (Deci & Ryan, 1994), teoria e atribuimit (Dweck, 2000), interesi (Hidi & Harackiewicz, 2000) dhe më e rëndësishme, të mësimi i vetë-rregulluar, përcaktuar nga Boekaerts (2006) si “një proces shumëplanësh, me shumë komponentë që synon efektin, njohjet dhe veprimet, si dhe karakteristikat e mjedisit për rregullim në shërbim të qëllimeve të dikujt” (f. 347). Ndërsa shumica e kërkimit të vetë-rregullimit ka prirjen t’i japë përparësi qasjeve njohëse ose motivuese, në vitet e fundit ka pasur disa përpjekje domethënëse për t’i afuar këto dy lidhje më afër, pasi siç argumenton Boekaerts (2006), të nxënësve i vetë-rregulluar udhëhiqet në mënyrë metakognitive dhe ngarkuar emocionalisht (f. 348).

Boekaerts ka propozuar një model të thjeshtë, por të fuqishëm për të kuptuar të nxënësve të vetë-rregulluar, i cili është quajtur teoria e përpunimit të dyfishtë (Boekaerts, 1993). Në model:

Supozohet se nxënësit të cilët janë ftuar të marrin pjesë në një aktivitet mësimor përdorin tre burime informacioni për të formuar një përfaqësim mendor të detyrës dhe për të vlerësuar atë: (1) perceptimet aktuale të detyrës dhe fizike, sociale dhe konteksti cilësor brenda të cilit është përcaktuar; (2) njohuri specifike të fushës dhe (meta) strategji njohëse që lidhen me detyrën; dhe (3) besime motivuese, duke përfshirë aftësitë specifike të fushës, besimet e interesit dhe përpjekjeve. (Boekaerts, 2006, f. 349)

Në varësi të rezultatit të vlerësimin, nxënës aktivizon vëmendjen përgjatë njëres nga dy rrugët: “rruga e rritjes” ku qëllimi është rritja e kompetencës ose “rruga e mirëqenies” ku vëmendja përqendrohet në parandalimin e kërcënimit, dëmit ose humbjes. Ndërsa opsioni i parë është dukshëm i preferueshëm, kjo e fundit nuk është domosdoshmërisht kundër-produktive – duke ndjekur rrugën e mirëqenies, nxënësi mund të gjejë një mënyrë për të rifituar mirëqenien (për shembull duke ulur koston e dështimit) që lejon zhvendosjen e energjisë dhe vëmendjen në rrugën e rritjes.

Nxënësit që janë personalisht të interesuar për një detyrë, padyshim ka të ngjarë të aktivizojnë energjinë gjatë rrugës së rritjes, por kur nxënësit nuk janë të interesuar personalisht për të detyrë, një numër i veçorive të detyrës mund të nxisë interes të situatës. Konsideratat e shkëmbimit ndërmjet vlerës së detyrës dhe koston do të ndikojë gjithashtu në atë që nxënësit i drejtojnë energjitë e tyre. Në veçanti, nxënësit ka gjasa të fokusohen në rritje sesa në mirëqenies kur shohin aftësinë si rritje sesa fikse (Dweck, 2000), kur kanë mjeshhtëri në vend të orientimit të performancës (Dweck, 2000) dhe kur identifikohen me qëllimin (Deci & Ryan, 1994).

Si përmbledhje, pasi të nxënit është i paparashikueshëm, vlerësimi është i nevojshëm për të bërë rregullimet përkatëse në udhëzime, por proceset e vlerësimit vetë ndikojnë në gatishmërinë, dëshirën dhe kapacitetin për të mësuar të nxënësit (Harlen & Deakin-Crick, 2002). Në mënyrë që vlerësimi të mbështesë të nxënit, ai duhet të sigurojë udhëzime mbi hapat e mëtejshëm të mësimin dhe duhet të ofrohet në një mënyrë që inkurajon nxënësin të drejtojë energjinë drejt rritjes, sesa mirëqënies.

6. Konkluzion

Ideja që vlerësimi mund të mbështesë mësimin nuk është një ide e re. Është e pakonceptueshme që personat e përfshirë në përpjekjet e hershme për të komunikuar idetë, aftësitë, ose praktikatat të të tjerët nuk e kuptuan që përpjekje të tilla nuk mund të garantoheshin të jenë të suksesshme dhe se udhëzimi efektiv kërkon vlerësim dhe përshtatje. Megjithatë, ka vetëm dyzet vjet që nga koha kur Benjamin Bloom sugjeroi për herë të parë se shqyrtimi i proceseve si vlerësimi mund të jenë të dobishme ose iluminuese. Në atë kohë, Bloom sugjeroi se procese të tilla mund të ishin më efektive të ndara nga përdorimi i vlerësimit për të regjistruar arrijtjet e nxënësve, por për njëzet vitet e ardhme, roli dominues i vlerësimit u shikuar si regjistrimi i arrijtjeve të nxënësit, megjithëse kishte një numër përpjekjesh për të përdorur provat e mbledhura me qëllim përmbledhjen e arrijtjes në mënyra të tjera, veçanërisht për të përmirësuar mësimin. Megjithatë, vetëm në fund të viteve '80 ideja se praktikatat e vlerësimit në klasë nuk mund të përballojnë dhe kufizojnë të nxënit e nxënësit filluan të pranohen në mënyrë të përhapur; vlerësimi i përdorur mund të përmirësonte në mënyrë thelbësore të nxënit, por shumicën e kohës, ndikimi i praktikave të vlerësimit ishte të limitonte, apo edhe të reduktonte të nxënit e nxënësit.

Gjatë viteve '90, një numër studimesh hulumtuan mbi idenë se vëmendja ndaj vlerësimit si pjesë integrale e mësimin mund të përmirësojë rezultatet e të nxënit për nxënësit, dhe në të njëjtën kohë, u bënë përpjekje për të lidhur praktikën në klasë me kërkime, veçanërisht reagim, motivim, atributim, dhe të mësuesit vetë-rregulluar. Për pjesën më të madhe të kohës, termi "vlerësim formues" nuk ishte përcaktuar saktësisht, dhe, si rezultat, studimet kërkimore mbi një aspekt të përdorimit të vlerësimit për të përmirësuar udhëzimet u përdorën si prova për të mbështetur efikasitetin e mjaft aspekteve pa lidhje. Pjesërisht, në përgjigje të kësaj, shumë autorë ndaluan së përdoruri termin "vlerësim formues" duke preferuar shprehjen "vlerësim për të nxënit", megjithëse përsëri kuptimi i tij i saktë rrallë përcaktohej, përtej idesë se vlerësimi duhet të përdoret gjatë udhëzimeve për të përmirësuar rezultatet e të nxënit.

Ky punim shqyrtoi këto zhvillime dhe përshkroi përpjekjet e fundit që janë bërë për të teorizuar vlerësimin formues dhe vlerësimin për të nxënit në një sërë mënyrash, konkretisht për sa i përket strategjive në klasë dhe teknikave praktike që mësuesit mund të përdorin për të përmirësuar cilësinë e provave mbi të cilat vendimet udhëzuese dhe nxënësit bëjnë. Ndërkohë që ka ende shumë punë për të bërë për të integruar kërkimin mbi vlerësimin për të nxënit me më shumë kërkime thelbësore mbi udhëzime hartim, reagim, të nxënit i vetërregulluar dhe motivim, tani ekziston një punë teorike dhe empirike e fuqishme që sugjeron se integrimi i vlerësimit me udhëzime mund të ketë fuqi të paparë për të rritur angazhimin e nxënësve dhe për të përmirësuar rezultatet e të mësimin.

Falënderim

Fig. 1, e cila u botua për herë të parë në Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Vlerësimi në klasa: Minutë pas minute dhe ditë pas dite. Drejtimi në arsim, 63(3), 18-24, është përdorur me lejen e Shërbimit të Testimit Arsimor.

Referenca

Allal, L., & Lopez, L. M. (2005). Formative assessment of learning: A review of publications in French [Vlerësimi formues i të nxënit: Shqyrtim botimesh në gjuhën frënge]. In J. Looney (Ed.), Formative assessment: Improving learning in secondary

- classrooms (fq.241-264). Paris, Francë: Organizata për Bashkëpunim dhe Zhvillim Ekonomik.
- Ausubel, D. P. (1968). Educational psychology: A cognitive view. New York, NY: Holt, Rinehart & Winston. [Ausubel, D. P. (1968). [Psikologjia edukative: Një shikim njohës, NY: Holt, Rinehart & Winston.]
- Bangert-Drowns, R. L., Kulik, C.-L.C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. Review of Educational Kërkim, 61(2), 213-238.
- [Bangert-Drowns, R. L., Kulik, C.-L.C., Kulik, J. A., & Morgan, M. (1991). Efektet udhëzues i reagimit në raste të ngjashme me testime. Rishikim i Kërkimit Arsimor 61(2), 213-238.
- Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C.-L.C. (1991) Effects of frequent classroom testing [Efektet e testimit të shpeshtë në klasë. Gazeta e Kërkimit Arsimor, 85(2), 89-99.
- Bennett, R. E. (2009). A critical look at the meaning and basis of formative assessment (ETS RM-09-06). Princeton, NJ: Educational Testing Service. [Një vështrim kritik mbi kuptimin dhe bazën e vlerësimit]
- Black, H. (1986). Assessment for learning. In D. L. Nuttall (Ed.), Assessing educational achievement (pp. 7-18). London: Falmer Press. [Vlerësimi për mësimin, Vlerësimi i arrijtjes mësimor]
- Black, P. J., & Wiliam, D. (1998a). Assessment and classroom learning. Assessment in Education: Principles, Policy and Practice, 5(1), 7-74. [Vlerësimi dhe mësimi në klasa. Vlerësimi në Arsim: Parimet, Politika dhe Praktika]
- Black, P. J., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan, 80(2), 139-148. [Brenda kutisë së zezë: Rritja e standardeve përmes vlerësimit në klasë]
- Black, P. J., & Wiliam, D. (2004). The formative purpose: assessment must first promote learning. In Wilson, M. (Ed.). Towards coherence between classroom assessment and accountability: 103rd Yearbook of the National Society for the Study of Education (part 2) (vol. Part II, pp. 20-50). Chicago, IL: University of Chicago Press [Qëllimi formues: vlerësimi duhet fillimisht të promovojë mësimin. Libri vjetor i Shoqërisë Kombëtare për Studimin e Edukimit (pjesa e dytë)].
- Black, P. J., & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability, 21(1), 5-31. [Zhvillimi i teorisë së vlerësimit formues. Vlerësimi Arsimor, Vlerësimi dhe Logjardhënia]
- Black, P., & Wiliam, D. (2011). Developing a theory of formative assessment. In J. Gardner (Ed.), Assessment and learning (2nd ed.). London, UK: Sage. [Zhvillimi i teorisë së vlerësimit formues, Vlerësimi dhe mësimi]
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. Phi Delta Kappan, 86(1), 8-21. [Duke punuar brenda kutisë së zezë: Vlerësimi për mësimin në klasë]
- Bloom, B. S. (1984a). The search for methods of instruction as effective as one-to-one tutoring. Educational Leadership, 41(8), 4-17. [Kërkimi i metodave të mësimdhënies po aq efektive sa mësimi individual. Udhëheqja arsimore]
- Bloom, B. S. (1984b). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13(6), 4-16. [Problemi 2-sigma: Kërkimi për metoda të mësimin në grup po aq efektive sa edhe mësimi individual. Kërkuesi Arsimor]
- Boekaerts, M. (1993). Being concerned with well being and with learning. Educational Psychologist, 28(2), 149-167. [Shqetësimi për mirëqenien dhe mësimin. Psikolog Arsimor]
- Boekaerts, M. (2006). Self-regulation and effort investment. In K. A. Renninger & I. E. Sigel (Eds.), Handbook of child psychology volume 4: Child psychology in practice (6th ed., pp. 345-377). New York, NY: Wiley. [Vetë-rregullimi dhe investimi në përpjekje. Në K. A. Renninger & I. E. Sigel (Eds.), Manual i psikologjisë së fëmijëve vëllimi 4: Psikologjia e fëmijëve në praktikë]
- Broadfoot, P. M., Daugherty, R., Gardner, J., Gipps, C. V., Harlen, W., James, M., et al. (1999). Assessment for learning: Beyond the black box. Cambridge, UK: University of Cambridge School of Education. [Vlerësimi për arsim: Përtej kutisë së zezë]
- Broadfoot, P. M., Daugherty, R., Gardner, J., Harlen, W., James, M., & Stobart, G. (2002). Assessment for learning: 10 principles. Cambridge, UK: University of Cambridge School of Education. [Vlerësimi për mësimin: 10 parime]
- Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. Teachers College Record, 106(3), 429-458. [Vlerësimi në klasa: Tensionet dhe kryqëzimet në teori dhe praktikë]
- Brookhart, S. M. (2007). Expanding views about formative classroom assessment: A review of the literature. In J. H. McMillan (Ed.), Formative classroom assessment: Theory into practice (pp. 43-62). New York, NY: Teachers College Press. [Zgjerimi i pikëpamjeve rreth vlerësimit formues të klasës: Një përmbledhje e literaturës. Në J. H. McMillan (Ed.), Vlerësimi formues i klasës: Teoria në praktikë]
- Brophy, J. (1981). Teacher praise: A functional analysis. Review of Educational Research, 51(1), 5-32. [Lavdërimi i mësuesit: Një analizë funksionale. Shqyrtimi i Kërkimit Arsimor]
- Carroll, L. (1871). Through the looking glass and what Alice found there. London, UK: Macmillan. [Nëpërmjet xhamit dhe çfarë gjeti Alice atje]
- Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. Assessment in Education: Principles, Policy and Practice, 6(1), 32-42. [Model i vlerësimit formues në edukimin shkencor]
- Crooks, T. J. (1988). The impact of classroom evaluation practices on nxënësit. Review of Educational Research, 58(4), 438-481. [Impakti i praktikave të vlerësimit mbi nxënësit. Shqyrtimi i Kërkimit Arsimor]
- Deci, E. L., & Ryan, R. M. (1994). Promoting self-determined education. Scandinavian Journal of Educational Research, 38(1), 3-14. [Promovimi i arsimit të vetë-përcaktuar.]
- Dempster, F. N. (1991). Synthesis of research on reviews and tests. Educational Leadership, 48(7), 71-76. [Sinteza e hulumtimit mbi rishikimet dhe provimet. Lidhshmi mësimor]
- Dempster, F. N. (1992). Using tests to promote learning: A neglected classroom resource. Journal of Research and Development in Education, 25(4), 213-217. [Përdorimi i testeve për të promovuar mësimin: Burim i një klase të neglizhuar.]
- Dweck, C. S. (2000). Self-theories: Their role in motivation, personality and development. Philadelphia, PA: Psychology Press. [Teoritë e vetes: Roli i tyre në motivim, personalitet dhe zhvillim]
- Elshout-Mohr, M. (1994). Feedback in self-instruction. European Education, 26(2), 58-73. [Reagimi në vetë-mësim. Edukim Evropian]
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation—A meta-

- analysis. *Exceptional Children*, 53(3), 199-208. [Efektet e vlerësimit formues sistematik – Meta-analizë. Fëmijë të veçantë]
- Guskey, T. R. (2010). *Formative assessment: The contributions of Benjamin S. Bloom*. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 106-124). New York, NY: Taylor & Francis. [Vlerësimi formues: Kontributet e Benjamin S. Bloom, Manual i vlerësimit formues]
- Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). (1998). *Metacognition in educational theory and practice*. Mahwah, NJ: Lawrence Erlbaum Associates. [Metakognjtim në teorinë dhe praktikën arsimore]
- Harlen, W., & Deakin-Crick, R. (2002). A systematic review of the impact of summative assessment and tests on nxënës' motivation for learning. In EPPI-Centre (Ed.), *Research evidence in education library* (1.1 ed., pp. 153-). London, UK: University of London Institute of Education Social Science Research Unit. [Një rishikim sistematik mbi impaktin e vlerësimit përmbledhës dhe testimetve mbi motivimin e nxënësëve për mësim]
- Hattie, J. (1999, August). *Influences on nxënës learning* Retrieved from: <http://www.education.auckland.ac.nz/uoa/education/staff/j.hattie/papers/influences.cfm> [Ndikimet mbi mësimin e nxënësëve]
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. [Fuqja e reagimit. Rishikim i Kërkimit Arsimor]
- Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, 70(2), 151-179. [Duke motivuar të pamotivuarit akademikë: Një çështje kritike për shekullin e 21-të]
- James, M. (1992, April). *Assessment for learning*. Paper presented at the Annual Conference of the Association for Supervision and Curriculum Development (Assembleja e Kritikës së Reformave të Vlerësimit dhe Testimit në Britani) held at New Orleans, LA. [Vlerësimi për mësimin. Punimi i prezantuar në Konferencën Vjetore të Shoqatës për Mbiqëqyrje dhe Zhvillim të Kurrikulave (seanca e Asamblesë me temë : Kritika e reformave mbi vlerësim dhe testimin në Britani) mbajtur në New Orleans, LA.]
- Kahl, S. (2005). Where in the world are formative tests? Right under your nose!. *Education Week*, 25(September (4)), 11. [Ku janë testimet formuese? Poshtë hundës! Java e arsimit, 25 shtator]
- Klenowski, V. (2009). Editorial: Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy, and Practice*, 16(3), 263-268. [Editorial: Vlerësimi për mësimin e rishikuar: Perspektiva në Azi-Pacifik]
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284. [Efektet e reagimit mbi performancën: Një rishikim historik, meta-analizë dhe reagim paraprak i teorisë së ndërhyrjes. Buletin Psikologjik]
- Koller, O. (2005). *Formative assessment in classrooms: A review of the empirical German literature*. In J. Looney (Ed.), *Formative assessment: Improving learning in secondary classrooms* (pp. 265-279). Paris, France: Organisation for Economic Cooperation and Development. [Vlerësimi formues në klasa: Një rishikim i literaturës empirike gjermane. Në J. Looney (Ed.), *Vlerësimi formues: Përmirësimi i mësimin në klasat e mesme* (f.265-279). Paris, Francë]
- Lewis, C. C. (2002). *Lesson study: A handbook of teacher-led instructional change*. Philadelphia, PA: Research for Better Schools. [Studim i mësimin: Një manual i ndryshimit të bazuar në udhëzime drejtuar nga mësues. Philadelphia, PA: Kërkim për Shkollat më të Mira]
- Looney, J. (Ed.). (2005). *Formative assessment: Improving learning in secondary classrooms*. Paris, France: Organisation for Economic Cooperation and Development. [Vlerësimi formues: Përmirësimi i mësimin në klasat e mesme. Paris, Francë: Organizata për Zhvillim dhe Bashkëpunim Ekonomik]
- Meisels, S. J., Atkins-Burnett, S., Xue, Y., Bickel, D. D., & Son, S.-H. (2003). *Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores*. *Education Policy Analysis Archives* 11(9). [Krijimi i një sistemi llogaridhënie: Impakti i vlerësimit mbi arritjet e fëmijëve në provime]
- Mitchell, R. (1992). *Testing for learning*. New York, USA: Free Press, Macmillan. [Testimi për të nxënët]
- Natriello, G. (1987). The impact of evaluation processes on nxënës. *Educational Psychologist*, 22(2), 155-175. [Impakti i procesve të vlerësimit mbi nxënësët]
- Nyquist, J. B. (2003). The benefits of reconstruing feedback as a larger system of assessment: A meta-analysis. Unpublished Master of Science thesis, Vanderbilt University. [Përfitimet e dhënies së reagimit si sistem më i madh i vlerësimit formues: Meta-analizë, Tezë e papublikuar Masteri shkencor]
- Parkhurst, H. (1922). *Education on the Dalton Plan*. London, UK: G Bell and Sons Ltd. [Arsimi në Planin Dalton]
- Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning. Towards a wider conceptual field. *Assessment in Education: Principles, Policy and Practice*, 5(1), 85-102. [Nga vlerësimi formues në rregullim të kontrolluar të mësimin. Drejt një fushe konceptuale më të gjerë. Vlerësimi në Arsim: Parimet, Politika dhe Praktika]
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science*, 28(1), 4-13. [Mbi përkufizimin e reagimit]
- Reiser, R. A. (1986). *Instructional technology: A history*. In R. M. Gagné (Ed.), *Instructional technology: Foundations* (pp. 11-48). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rheinberg, F. (1980). *Leistungsbewertung und lernmotivation. Achievement evaluation and learning motivation*. Gottingen, Germany: Hogrefe. [Vlerësimi i arritjeve dhe motivimi për mësim]
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144. [Vlerësimi formues dhe hartimi i sistemeve udhëzuese]
- Shepard, L. A., Hammerness, K., Darling-Hammond, L., Rust, F., Snowden, J. B., Gordon, E., et al. (2005). *Assessment*. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 275-326). San Francisco, CA: Jossey-Bass. [Vlerësim. Përgatitja e mësuesve për një botë në ndryshim]
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189. [Fokus mbi reagimin formues]
- Stiggins, R. J. (2005). From formative assessment to assessment FOR learning: A path to success in standards-based schools. *Phi Delta Kappan*, 87(4), 324-328. [Nga vlerësimi formues në vlerësim PER mësim: Rruga e suksesit në shkollat me standard bazë]
- Sutton, R. (1995). *Assessment for learning*. Salford, UK: RS Publications. [Vlerësimi për mësimin]
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment*. Buckingham, UK: Open University Press. [Hetimi i vlerësimit formues]
- Tunstall, P., & Gipps, C. (1996a). 'How does your teacher help you to make your work better?' Children's understanding of formative assessment. *The Curriculum Journal*, 7(2), 185-203. [Si mund t'iu ndihmojë mësuesi në përmirësimin e punës tuaj?]
- Tunstall, P., & Gipps, C. V. (1996b). *Teacher feedback to young children in formative assessment: A typology*. *British Educational Research Journal*, 22(4), 389-404. [reagimi i mësuesit për fëmijë në vlerësimin formues: Tipologji.]
- Vinner, S. (1997). From intuition to inhibition—Mathematics, education and other endangered species. In E. Pehkonen (Ed.), *In Proceedings of the 21st conference of the international group for the psychology of mathematics education*, Vol. 1 (pp. 63-78). Lahti, Finland: University of Helsinki Lahti Research and Training Centre. [Nga intuicioni në ndalim – Matematika, arsimit dhe specitë tjera në rrezik. Në E. Pehkonen (Ed.), *Ne procedimet e konferencës së 21-të të grupit ndërkombëtar për psikologjinë e mësimin të matematikës*]
- Wiener, N. (1948). *Cybernetics, or control and communication in the animal and the machine*. New York, NY: John Wiley & Sons Inc. [Kibernetika, ose kontrolli dhe komunikimi në kafshë dhe makineri]
- Wiliam, D. (2007). *Keeping learning on track: Classroom assessment and the regulation of learning*. In F. K. Lester, Jr., (Ed.). *Second handbook of mathematics teaching and learning* (pp. 1053-1098). Greenwich, CT: Information Age Publishing. [Mbajtja në shina e arsimit: Vlerësimi në klasa dhe rregullorë e mësimin. Në F. K. Lester, Jr., (Ed.). *Manuali i dytë i mësimin dhe studimit të matematikës*]
- Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18-40). New York, NY: Taylor & Francis. [Një përmbledhje e integruar e literaturës kërkimore dhe implikimet për teorinë e re të vlerësimit formues]
- Wiliam, D., & Thompson, M. (2008). *Integrating assessment with instruction: What will it take to make it work?* In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). Mahwah, NJ: Lawrence Erlbaum Associates. [Integrimi i vlerësimit me udhëzime: Çfare duhet për ta bërë funksionale? Në C. A. Dwyer (Ed.), *E ardhmja e vlerësimit: Formimi i mësimdhënies dhe mësimin* (f. 53-82)]
- Wiliam, D., Lee, C., Harrison, C., & Black, P. J. (2004). *Teachers developing assessment for learning: Impact on nxënës achievement*. *Assessment in Education: Principles, Policy and Practice*, 11(1), 49-65. [Mësuesit që zhvillojnë vlerësimin për të nxënët: Ndikimi në arritjet e nxënësit. Vlerësimi në arsim: Parimet, politika dhe Praktika]