Assessment as Learning: naar een nieuwe beoordelingscultuur.

Auteurs:

Simla Arikan, Assoc. Prof. Dr.
Wibran Dochy, Drs.
Filip Dochy, Prof. Dr. (1% top researcher USERN)
Mien Segers, Prof. Dr. (Francqui-prize winner 2021)

Contact: W. Dochy
Info@highimpactlearningthatlasts.com

‘Assessment-as-Learning’ als een van de bouwstenen van High Impact Learning (Dochy, Segers & Dochy, 2020) kan je lezen als ‘het verzamelen en interpreteren van betekenisvolle informatie over het leerproces en de leeropbrengsten als integraal onderdeel van het leerproces’. De kern is ‘feedback’, en daarmee maakt Assessment-as-Learning komaf met de ‘testcultuur` (Birenbaum, 1996).

Harland en Wald (2021) spreken over de “arms race” als metafoor om het fenomeen van het (her en der nog steeds gebruikte) grote aantallen toetsen te verklaren.
De metafoor van de 'wapenwedloop' werd opzettelijk gebruikt omdat het een proces van beoordelingsproliferatie illustreert, waarbij, volgens dit onderzoek, docenten drie specifieke vooronderstellingen hebben over de motivatie van lerenden: (1) Lerenden werken niet voor iets dat geen cijfer oplevert, en dus geven leraren geen opdrachten waar geen cijfer aan vasthangt; (2) in het modulaire systeem krijgen lerenden beperkte keuze en lesgevers concurreren voortdurend met elkaar om de tijd en aandacht van lerende, waarbij men dus cijfers moet inzetten als beloning om de concurrentie aan te kunnen; (3) lerenden zijn medeplichtig aan de wapenwedloop omdat ze aandringen op veel kleine toetsen als een soort vangnet. In een dergelijke testcultuur (Birenbaum, 1996) telt elk cijfer en ook cijfers op kleine (zogenaamde formatieve) toetsen worden gezien als 'high stakes' omdat ze allen veelal bijdragen tot een eindcijfer (Harland & Wald, 2021).

Toetsresultaten zouden informatie geven over waar de lerende staat in de verwerving van beoogde competenties of het behalen van de doelen. Ze bieden dus de mogelijkheid tot zogenoemde ‘formatieve feedback’ en zijn tegelijk input voor docenten om risico-studenten te identificeren.
Om van deze voordelen te kunnen genieten is een cijfer geven dat meetelt echter overbodig. Leeractiviteiten zoals opdrachten, essays, etc. bieden volgens Harland en Wald (2021) een rijkheid aan informatie over leerproces en leeropbrengsten en zijn dus een relevante bron voor feedback en het tijdig identificeren van studenten met problemen zodat deze tijdig hulp kunnen inroepen. Daarenboven heeft een ‘cijferarme’ opleiding vele andere voordelen. Bij veelvuldige toetsen met cijfers heeft een lerende geen keuze om al dan niet betrokken te zijn, of de vrijheid om naar eigen inzicht deel te nemen. Hij kan dus ook geen redelijke agency uitoefenen en neemt studiebeslissingen die rationeel zijn in termen van cijfers inclusief het kiezen van modules omdat ze ‘makkelijk’ zijn (zelfs als ze er geen interesse in hebben), omdat ze het juiste aantal credit points opleveren, omdat ze geen inspanning vergen, etc. Harland spreekt van een onderwerping aan de eisen van het beoordelingsregime.
De wapenwedloop/ testcultuur is een didactiek van controle (Wass et al. 2015) die het vermogen ontneemt om zelfstandig te studeren. De continue dreiging van een summatieve test is een machtsspel en leidt volgens Harland en Wald (2021) tot ondergeschiktheid en domesticering van de studenten en creëert een afhankelijkheidscultuur met controleerbare of 'tamme' leerlingen. In een cijfergedreven systeem is geen ruimte voor diepgang en essentiële interactie, zelfsturing of engagement.

Waarom houden opleidingen dan zo strak vast aan de toets- en cijfercultuur? Vele opleiders twijfelen er sterk aan of lerenden nog inspanningen zullen leveren als er geen cijfers te verdienen zijn. Ecclestone betoogde reeds in 1999 dat een dergelijke instrumentele houding van gewoon doen wat nodig is voor de toetsen of enkel doen wat moet, ervoor zorgt dat lerenden niet langer bereid zijn iets uitdagends te doen of iets cognitief complex uit te zoeken of op te lossen. Dit heeft alles te maken met twijfels over 'urgentie', 'gedrevenheid van lerenden', 'learner agency' en of lerenden wel verantwoordelijkheid zullen opnemen. De vele 'High Impact Learning' opleidingen in vele Europese landen tonen duidelijk een ander beeld: lerenden zetten zich in, nemen verantwoordelijkheid (zowel twaalfjarigen als twintigjarigen) en werken harder met meer plezier (Dochy et al., 2020). Kortom: er zijn alternatieven die werken.

Ook na de opkomst van de ideeën rond 3G assessment ging de wetenschap verder in die richting. Dylan Wiliam schreef in 2013: "… there is now a strong body of theoretical and empirical work that suggests that integrating assessment with instruction may well have unprecedented power to increase student engagement and to improve learning outcome" (Wiliam, 2011,p. 13). We kunnen tien jaar laten alleen maar zeggen dat de evidentie waar Wiliam naar verwijst alleen maar is toegenomen.
Niemand twijfelt nog een het gegeven dat lerenden dwingen door een lange toetsbatterij voor te leggen, niet alleen demotiverend is maar dat er ook weinig impact van het leren is op lange termijn (zie voor een overzicht Dochy & Segers, 2018). Zoals Wiliam aangeeft is het een integratie van leren en beoordelen of 'assessment as learning' die een ongeziene kracht genereert die leidt tot engagement van de lerende en een verbetering van leerresultaten.

Assessment as learning

Assessment as learning is een vorm van programmagericht beoordelen. Binnen het nieuwe concept 'Assessment as Learning' vormen feedbackdialogen de ruggegraat. Wij conceptualiseren 'Assessment as Learning' als een continu assessment dat in het leerproces zit verweven en bestaat uit 'resource assessment' en 'agency assessment'.

Resource assessment

Resource assessment zorgt ervoor dat lerenden de basiskennis en vaardigheden en generieke competenties diepgaand leren en leeruitkomsten op een intermediair abstractieniveau kunnen gebruiken om door middel van rubrics hun vordering in de ontwikkeling te kunnen aantonen. Vaak wordt hier een portfolio-assessment gebruikt; soms te vaak. En niet zonder reden. Winstone vond meerdere studies waarin studenten en docenten aangeven dat feedbackportfolio's de onafhankelijkheid van lerenden bevorderen, reflectie en het stellen van doelen bevorderen, dialoog met coaches faciliteren en het zoeken naar feedback stimuleren (Winstone et al., 2017). Eigenlijk wat we zowat willen, maar variatie is erg noodzakelijk om impact van leerprocessen op peil te houden.
Om die reden citeerden we ook nog tien andere methodes die resource assessment met meerdere beoordelaars kunnen ondersteunen:

1. Portfolio assessments (met verantwoordingsverslag & eindgesprek)
2. Two-stage assessments
3. Producten als input voor beoordeling
4. Procesverslag als input voor assessment
5. 180° feedback
6. 360° feedback
7. Criteriumgericht interview (bepaal criteria in interactie met de lerenden)
8. Peer-& co-assessment
9. Assessmentgesprekken in team (waarin nieuwe kennis geconstrueerd)
10. Assessmentgesprek/ Eindgesprek individueel
11. Exit demonstraties/presentaties*

Agency assessment

Elke beoordelingsmethode heeft zijn voor- en nadelen en lerenden proberen vaak te anticiperen op wat in een beoordeling wordt bekeken. Ideaal is uiteraard dat lerenden op geen moment denken aan een beoordeling omdat de integratie van leren en beoordelen dan zo sterk is dat het natuurlijk aanvoelt omdat men gedreven is om te leren. Die situatie maakten we zelf al vaker mee en studenten waren enthousiaste getuigen hiervan. Toch kan over cohorten heen ook datgene waar beoordelaars op letten gaan fungeren als signalen van 'wat belangrijk is'. Om die reden stelde James (2012) voor om ook 'agency assessment' expliciet te maken. Volgens James is 'agency assesment' *het beoordelen van de mate van 'agency' (beheer) waarmee men know-how en tools inzet voor probleemanalyse, oplossing, en productgericht werken en de eigen evaluatie van/ reflectie op dit proces*. De borging die men doet in de vele feedbackverslagen die lerenden maken op basis van de cyclische feedbackdialogen geven veelal voldoende informatie om hier een oordeel over te geven, mogelijks door meerdere beoordelaars.

In een leerproces van bijvoorbeeld een semester, wat ongeveer 20 weken zou kunnen omvatten (omdat er ook geen studieweken of toetsweken nodig zijn) zou men bijvoorbeeld 10 à 20 feedbackdialogen kunnen hebben, 1 eindproduct en 2 tussenproducten, een procesverslag, een 180° feedback, een CGI en 2 peer-assesments. Dat zouden bijvoorbeeld 18 datapunten zijn voor een beoordeling van dit semester. Dat aantal datapunten moet een goed en betrouwbaar beeld geven van de behaalde leeruitkomsten met een sterke kennisborging.

Meer over AaL leest u in het 2022 boek 'Feedback met impact is dialoog' en onze HILL inspirational paper over AaL als programmagericht of curriculumgericht beoordelen.

* zie ook: Dochy, F. & Segers, M. (2022). Feedback met impact - is dialoog. Kern van VIP-Coaching en Assessment-as-Learning. Amsterdam: Boom. https://www.boomhogeronderwijs.nl/actueel-item/80-3733_Feedback-met-impact-is-dialoog

Dochy, F. & Segers, M. (2022). Feedback met impact - is dialoog. Kern van VIP-Coaching en Assessment-as-Learning. Amsterdam: Boom.
https://www.boomhogeronderwijs.nl/actueel-item/80-3733_Feedback-met-impact-is-dialoog

Dochy, F. & Segers, M. (2020).Bouwstenen voor High Impact Learning: Van model naar tips en tricks en succescriteria.
https://www.boomhogeronderwijs.nl/product/100-9221_Bouwstenen-voor-High-Impact-Learning-2e-druk

Dochy, F. & Segers, M. (2018). Creating impact through future learning: The High Impact Learning that Lasts Model. London: Routledge Publishers.
https://www.routledge.com/Creating-Impact-Through-Future-Learning-The-High-Impact-Learning-that/Dochy-Segers/p/book/9781138577886 (20% DISCOUNT with code FLR40 when ordering via web).

Behoefte aan keynote, worklab, expert support? Info@highimpactlearningthatlasts.com