



Het oog van de meester en transparantie voor de student gaan goed samen

Een aanpak die holistisch beoordelen en een helder leerproces voor de student verenigt

Analyse | door Marion Tillema

17 maart 2020 | Het is mogelijk om helder te zijn over eisen voor een opdracht richting de student, zonder in dichtgetimmerde beoordelingsformulieren te vervallen. Volgens Marion Tillema (Avans) ligt de sleutel voor het dilemma in het toepassen van een holistische benadering, bijvoorbeeld in een comparatieve beoordeling.



ScienceGuide is bij wet verplicht je toestemming te vragen voor het gebruik van cookies. Lees [hier](#) over ons cookiebeleid en klik op OK om akkoord te gaan

OK



Recente bijdragen van [Menno van der Schoot](#) en [Alex Steenbreker](#) over de waarde van *rubrics* werpen licht op de veronderstelde nadelen en voordelen van deze analytische beoordelingsmethode voor het beoordelen van complex studentwerk. Van der Schoot benadrukt vooral het validiteitsprobleem van analytische oordelen en vraagt om ruimte voor de ervaring van de docent in de beoordeling. Steenbreker pleit voor transparante eisen voor beoordeling voor studenten.

Alhoewel deze twee voorwaarden elkaar op het eerste gezicht enigszins lijken te bijten is het een misvatting te denken dat deze zaken niet te verenigen zijn. Laten we eens kijken naar die misvattingen, en een voorbeeld geven van een aanpak die validiteit, een plaats voor de expertise van de docent en voldoende transparantie voor de student verenigt.

Het is niet goed of fout, het is complex

Complexe studentproducten, zoals een onderzoeksthese in het wo of adviezen, technische constructies of ontwerpen in het hbo, hebben twee kenmerken. Ten eerste gaat het om producten die niet simpelweg goed of fout zijn. Het bepalen van de kwaliteit ervan is geen kwestie van scores, maar van beoordelen, waarbij interpretatie – door de docent – een rol speelt.

Daarnaast wordt de kwaliteit ervan bepaald door meerdere aspecten, en is die kwaliteit bovendien eerder een interactie van die aspecten dan een optelsom ervan. Per werk kunnen andere (combinaties van) aspecten saillant zijn. Hieraan refereert Menno van der Schoot als hij schrijft dat de verschillende onderdelen van een creatief-intellectuele prestatie op elkaar inwerken en dat het onmogelijk lijkt om te doorgronden hoe die wisselwerkingen de gehele kwaliteit van dergelijke prestaties bepalen (zie ook Lesterhuis et al., 2015). Dit betekent dat een vastgelegde weging van aspecten, zoals in rubrics, per definitie niet passend is.

””Een vastgelegde weging van aspecten is per definitie niet passend voor complexe producten””



Van der Schoot doet in de conclusie van zijn stuk dan ook de suggestie om meer holistische beoordelingsmethoden te hanteren, en noemt comparatief beoordelen daarbij in het bijzonder. Hij noemt als argument hiervoor dat holistische

Meest Gelezen



Universiteiten zwijgen als hoogleraren zich vergrijpen



Marcel Levi probeert onrust binnen NWO te sussen



'Over acht jaar moet hoger onderwijs bijna overal hetzelfde georganiseerd zijn'



Rente op studieschuld wordt 0,46 procent



Terugloop aantal mbo-studenten zeer zorgelijk

Bij inschrijving ga je akkoord met onze privacy-voorwaarden. Deze voorwaarden zijn [hier](#) te lezen.

beoordelingsmethoden niet minder betrouwbaar zijn (of in het geval van comparatief beoordelen: een hogere betrouwbaarheid kennen) dan analytische beoordelingsmethoden. Analytische oordelen zijn in deze context oordelen waarbij het totaaloordeel een optelsom is van deeloordelen op verschillende aspecten, waarbij de weging van deze aspecten van tevoren wordt vastgesteld en voor elk te beoordelen werk hetzelfde is. Holistische oordelen zijn oordelen waarbij de invloed van deelaspecten op het totaal niet vastligt.

Holistisch beoordelen voor validiteit

Holistisch beoordelen kent inderdaad voordelen. Holistisch beoordelen doet recht aan complex werk, omdat het ruimte laat aan de eerdergenoemde interactie tussen verschillende kwaliteitsaspecten van het te beoordelen product . Daarnaast kennen holistische oordelen vermoedelijk een hogere generaliseerbaarheid: ze zijn betere voorspellers van soortgelijke prestaties op een latere taak dan analytische oordelen, die meer taakafhankelijk lijken te zijn. Schoonen (2005) en Van den Bergh et al. (2012) toonden dit bijvoorbeeld aan voor het beoordelen van schrijfvaardigheid.

Bij analytische oordelen zijn de resulterende scores of cijfers vaak meer specifiek voor het onderwerp waarover geschreven wordt. Moeten studenten op een later moment iets schrijven over een ander onderwerp, dan is het cijfer op de vorige opdracht een minder sterke voorspeller van de volgende prestatie dan bij holistische oordelen het geval is.

Mogelijk zijn analytische beoordelingsschema's zoals *rubrics* in hun precieze omschrijvingen van de gewenste prestaties meer toegespitst op de specifieke opdracht en minder op de vaardigheid in het algemeen. Generaliseerbaarheid is natuurlijk een belangrijk criterium voor een beoordeling: we willen uiteindelijk weten of een student klaar is voor uitoefening van zijn of haar vak.

Daarnaast vinden docenten holistisch beoordelen in het algemeen bevredigender: men vindt dat de beoordelingen meer recht doen aan het werk en de aanpak – we werken in de praktijk van mijn eigen opleiding met holistisch paarsgewijs vergelijken – werkt prettiger dan scoren met behulp van rubrics. Dit is op zichzelf natuurlijk geen reden voor het kiezen van deze beoordelingswijze, maar mooi meegenomen bij gelijke of betere validiteit.

Transparante eisen voor studenten

Rubrics verschaffen duidelijkheid, maar niet op de beste manier om studenten transparantie over de eisen te verschaffen. Alex Steenbreker benadrukt in haar reactie op Van der Schoot het belang van helderheid voor studenten over de eisen die aan het werk gesteld zullen worden. Het is inderdaad van groot belang dat beoordelingen van studentwerk naast betrouwbaar en valide ook transparant en navolgbaar zijn, en dat studenten van tevoren een goed beeld krijgen van de verwachtingen. Er zijn echter betere manieren om transparantie over de eisen te bereiken dan met rubrics, die bovendien goed toepasbaar zijn op holistisch beoordelen.

”Dat docenten holistisch beoordelen bevredigender vinden is op zichzelf geen reden, maar wel mooi meegenomen bij gelijke of betere validiteit”



Wat we in het onderwijs voorafgaand aan een beoordeling willen bereiken is niet alleen dat studenten de eisen aan de beoordeling kennen. We willen dat ze een goede representatie vormen van wat kwaliteit inhoudt voor het betreffende complexe producttype: wat is het kwaliteitsspectrum van een goede onderzoekstheze, van een goed adviesrapport, van een ontwerp?

Een *good practice* uit mijn eigen opleiding is om met studenten actief in gesprek te gaan op basis van voorbeeldwerk, zowel door de docent ingebracht voorbeeldwerk als tussentijds werk van de studenten zelf. De docent kan hierbij de voorbeelden toelichten: op basis van welke (interactie van) kenmerken is dit een goed werk?

Rijke input in de begeleiding

Vaak denken docenten dat feedback geven veel tijd vraagt in de voorbereiding van lessen. Feedback kan echter ook een constructief onderdeel van een actieve les met betrokken studenten zijn. Een goede werkvorm hiervoor is om studenten de werken te laten rangordenen (met digitale hulpmiddelen of fysiek in de ruimte) en de ordening te laten toelichten. De docent begeleidt dit gesprek om opvattingen, en ook misconcepties, over kwaliteit aan het licht te laten komen. Deze praktijk maakt gebruik van de principes van *modeling* (Bandura, 1975) en formatief handelen (Sluijsmans, 2019).

Mits goed begeleid door de docent levert deze aanpak veel rijkere input op voor studenten dan enkel eisen communiceren: er ontstaat een genuanceerde dialoog met studenten waarin de complexe aard van dit type werk beter tot uiting komt. Naast abstracte beschrijvingen van kwaliteitskenmerken levert deze aanpak studenten ook een beeld op van het bereik van mogelijkheden. Het loont dus om de tijdsinvestering die nodig is voor de constructie van een rubric in plaats daarvan te steken in transparantiebevorderend onderwijs, waarin een hoog niveau van begeleiding wordt bereikt.

Holistisch en betrouwbaar

Het bezwaar zou kunnen worden gemaakt dat holistische beoordelingswijzen willekeur in de hand kunnen werken. Juist door de ruimte die vrijkomt voor het expertoordeel over de wisselwerking tussen aspecten lopen we het risico dat er beoordelaarsfouten in de beoordeling sluipen. Diverse studies tonen echter aan dat holistische oordelen onder bepaalde voorwaarden voldoende betrouwbaar kunnen worden uitgevoerd.

““Feedback kan een constructief onderdeel van een actieve les met betrokken studenten zijn”“



Ten eerste is het belangrijk dat de beoordeling door meerdere, onafhankelijk opererende beoordelaars wordt uitgevoerd (Tillema et al., 2013; Van den Bergh et al., 2012). Ten tweede stelt een betrouwbare holistische beoordeling eisen aan de expertise, ervaring (training) en samenstelling van het beoordelaarsteam (Barkaoui, 2011; Harsch & Martin, 2013). Ten slotte draagt de beschikbaarheid van een ijkpunt bij aan de betrouwbaarheid van de beoordeling.

Ijpunten kunnen worden gecreëerd in de vorm van zorgvuldig geselecteerde voorbeeldwerken die punten op de kwaliteitschaal representeren, bijvoorbeeld: slecht, onvoldoende, gemiddeld, voldoende, goed. Beoordelaars moeten te beoordelen werken dan ergens op deze schaal plaatsen. Ook paarsgewijs vergelijken (Coertjens et al., 2015; Lesterhuis et al., 2015; Verhavert et al. 2018) biedt beoordelaars een ijkpunt en levert een betrouwbare beoordeling op: de beoordelaar moet per paar aangeven welk van de twee werken het betere is.

Kortom, het is mogelijk om een valide beoordeling van complex werk te verenigen met recht doen aan de expertise van de docent, voldoende transparantie voor de student en optimaal leren. Hiervoor is een slimme, geïnformeerde combinatie van maatregelen nodig. Hierbij kunnen holistische beoordelingswijzen onder specifieke voorwaarden worden ingezet. De beoordeling dient bovendien te worden voorafgegaan door onderwijs waarin studenten begeleid worden om te komen tot een goede representatie van wat kwaliteit inhoudt voor het betreffende complexe producttype.

feedback studiebegeleiding



Marion Tillema : Onderzoeker Brein & Leren

Marion Tillema is onderzoeker bij het lectoraat Brein & Leren (Avans Hogeschool), opleidingscoördinator Communication & Multimedia Design (Avans Hogeschool, locatie 's-Hertogenbosch) en bestuurslid van het ComeniusNetwerk. Als onderzoeker houdt ze zich bezig met de onderwerpen beoordelen en kritisch denken.



Literatuurverwijzingen

Analysis of modeling processes.

Bandura, A. (1975). *School Psychology Review* 4 (1), 4-10.

Effects of marking method and rater experience on ESL essay scores and rater performance.

Barkaoui, K. (2011). *Assessment in Education: Principles, Policy & Practice* 18 (3), 279-293.

Teksten beoordelen met criterialijsten of via paarsgewijze vergelijking: een afweging van betrouwbaarheid en tijdsinvestering.

Coertjens, L., Lesterhuis, M., Verhavert, S., Van Gasse, R. & De Maeyer, S. (2017). *Pedagogische studien* 94 (4), 283-303.

Comparing holistic and analytic scoring methods: issues of validity and reliability.

Harsch, C. & Martin, G. (2013). *Assessment in Education: Principles, Policy & Practice* 20 (3), 281-307.

Competenties kwaliteitsvol beoordelen: brengt een comparatieve aanpak soelaas?

Lesterhuis, M., Donche, V., De Maeyer, S., Van Daal, T., Van Gasse, R., Coertjens, L., Verhavert, S., Mortier, A., Coenen, T., Vlerick, P., Vanhoof, J. & Van Petegem, P. (2015). *Tijdschrift voor hoger onderwijs* 33 (2), 55-67.

Generalizability of writing scores: An application of structural equation modeling.

Schoonen, R. (2005). *Language Testing* 22 (1), 1-30.

Didactisch en formatief handelen: twee zijden van dezelfde medaille.

Sluijsmans, D. M. A. (2019). *Tijdschrift Hoger Onderwijs Management, novembernummer*, 4-7.

Quantifying the quality difference between L1 and L2 essays: A rating procedure with bilingual raters and L1 and L2 benchmark essays.

Tillema, M., van den Bergh, H., Rijlaarsdam, G. & Sanders, T. (2013). *Language Testing* 30 (1), 71-97.

Measuring writing: Recent insights into theory, methodology and practice.

Van den Bergh, H., De Maeyer, S., Van Weijen, D. & Tillema, M. (2012). Generalizability of text quality scores. In Van Steendam, E., Tillema, M., Rijlaarsdam, G.C.W. & Van den Bergh, H. (Eds.). (pp. 23-32). Brill.

A meta-analysis on the reliability of comparative judgement.

Verhavert, S., Bouwer, R., Donche, V., & De Maeyer, S. (2019). *Assessment in Education: Principles, Policy & Practice* 26 (5), 541-562.

ScienceGuide

Gerelateerde artikelen



Duurzaam onderwijs en onderzoek

Het kabinet wil weten hoe onderzoek en onderwijs de duurzame ontwikkeling...



Meester en juf academisch geschoold

Studenten kunnen voortaan gelijktijdig hun bachelordiploma onderwijskunde én hun lesbevoegdheid...



Inholland-student helpt popdiva

Hoeveel studenten kunnen dát op hun cv zetten? Inholland...

Schrijf je in voor onze nieuwsbrief

Klik hier



Laatste artikelen

'Academische vrijheid alleen geborgd met bedrijfsleven en overheid op afstand'

In EU-wetgeving wordt academische vrijheid vooral commercieel en niet...

Dijkgraaf bekritiseert studenten en vindt dat ze onterecht klagen

Het hoogstopgeleide deel van de bevolking krijgt veel en...

'Universiteiten mogen vaker hogeschooltje spelen'

Universiteiten kunnen nog veel leren van hogescholen als het...

'Praktijkgericht onderzoek is vaak al wetenschapscommunicatie'

Valorisatie is niet de kers, het is bij praktijkgericht...

'Open Education verbetert het onderwijs, maar het moet duurzamer'

Voor Robert Schuur, lector Open Education Resources bij Fontys...

Categorieën

HOGER ONDERWIJS
WETENSCHAP
POLITIEK & DEBAT
CARRIÈRE
INTERNATIONAAL
OPINIE

VOLG SCIENCEGUIDE OP



ScienceGuide BV

Schalkwijkstraat 10
3512 KS Utrecht

U kunt de redactie bereiken op

030-20 72 456
redactie@scienceguide.nl
scienceguide.nl
KvK-nummer 32119189