



**Holistisch beoordelen
vergroot de
pro-activiteit en
creativiteit voor
studenten**

Holistisch beoordelen van meerdere samenhangende leeruitkomsten bij flexibel deeltijdonderwijs

■ Anne Lohuis

A. Lohuis MSc is onderwijskundige en onderzoeker bij Saxion Parttime School bij Saxion Hogeschool. E-mail: a.h.lohuis@saxion.nl

Saxion Parttime School (SPS) (zie Kader 1) biedt flexibel deeltijdonderwijs aan waarin de dagelijkse praktijk van studenten centraal staat. In voorgaande jaren lag de focus op het apart beoordelen van steeds één leeruitkomst met een voorgeschreven beroepsproduct. Om meer flexibiliteit te realiseren en met het product een grotere impact te bereiken in de praktijk is het nu mogelijk dat studenten samenhangende leeruitkomsten in een toets kunnen combineren en zelf producten kiezen om deze leeruitkomsten aan te tonen. Deze samenhangende leeruitkomsten worden holistisch beoordeeld. Waarom sluit juist holistisch beoordelen hier goed bij aan en waarin verschilt dit van analytisch beoordelen? Hoe ziet de praktische toepassing van holistisch beoordelen van samenhangende leeruitkomsten eruit? In dit artikel zullen deze vragen worden beantwoord.

Saxion Parttime School

Saxion Parttime School (SPS) is de academie voor het deeltijdonderwijs van Saxion Hogeschool. SPS streeft er naar om onderwijs aan te bieden dat aansluit bij de behoeften van zowel bedrijven, werkenden als werkzoekenden. Het onderwijs is flexibel in tijd, plaats, tempo en momenten van instroom voor studenten. Daarnaast is het onderwijs blended wat inhoudt dat er een mix is van bijeenkomsten op locatie, online leren en leren op de werkplek. Vraagstukken uit de eigen praktijk staan centraal zowel in het onderwijs als toetsing zodat opgedane kennis en vaardigheden direct in de dagelijkse praktijk kunnen worden toegepast. In het curriculum van een opleiding zitten verplichte sectormodulen (modulen die gevolgd worden door alle opleidingen in een sector), clustermodulen (modulen die gevolgd worden door tenminste twee opleidingen) en opleidingsmodulen (modulen die gelden voor één opleiding).

Holistisch beoordelen versus analytisch beoordelen

Wanneer het gaat om manieren van beoordelen, wordt vaak het strikte onderscheid gemaakt tussen het gebruik van een holistische of analytische beoordelingsschaal. Bij een zuiver analytische manier van beoordelen is een beoordelingsinstrument opgedeeld in meerdere criteria die afzonderlijk van elkaar beoordeeld worden. Deze scores worden later bij elkaar opgeteld waarna er een eindoordeel wordt vastgesteld (Hunter et al., 1996; Nitko, 2001 & Mertler, 2000). Bij een zuiver holistische manier van beoordelen is de kwaliteit van de gehele prestatie of het product het uitgangspunt zonder dat specifieke deelaspecten worden gescoord (Barylá et al., 2012; Jonsson & Svingby, 2007; Nitko, 2001; Sadler, 2009; Schunn et al., 2016). Het verschil tussen deze twee wordt daarmee met name zichtbaar in het aantal beoordelingscriteria waarmee een prestatie of product wordt beoordeeld en de manier waarop men tot het eindoordeel komt. Het onderscheid tussen holistisch en analytisch beoordelen blijkt echter niet zo zwart-wit. Je zou het kunnen plaatsen op een continuüm waarin verschillende varianten zijn te onderscheiden die onder een holistische dan wel analytische benadering vallen. Een voorbeeld van zo'n continuüm met meerdere varianten van de twee benaderingen is die van Hunter et al. (1996) (Figuur 1). Zij onderscheiden vijf varianten.

De verschillen tussen deze varianten zitten met name in het aantal criteria, de totstandkoming van het oordeel en de feedback naar studenten. Hoe meer naar rechts op het continuüm, hoe gedetailleerder de beoordeling wordt. Bij 'analytic scoring' beoordeel je criteria die je terug wilt zien in het product of de prestatie van de student afzonderlijk. Het optellen van de score op deze afzonderlijke criteria zorgt voor het totaaloordeel. Bij 'atomistic scoring' ga je nog een stap verder en kijk je bijvoorbeeld exact naar de aan- of afwezigheid van bepaalde fouten, zoals het aantal spelfouten in een verslag (Cooper, 1977; Goulden, 1992; Hunter et al, 1996; Lloyd-Jones, 1977). De meest linkse en daarmee meest holistische benadering, 'general impression scoring', kent geen vastgestelde criteria die men in de prestatie of het product van de student terug wil zien (Goulden, 1992; Hunter et al., 1996). Na het lezen van bijvoorbeeld een verslag wordt door de beoordelaar in een keer een score toegekend op basis van de eigen professionele norm, echter wel binnen de context waarvoor de beoordelaar wordt ingezet. Daarmee gaat deze benadering nog een stap verder dan 'holistic scoring' omdat daarbij wel vooraf, op basis van een gemeenschappelijke norm, enkele en meer algemene criteria zijn bepaald om het algehele oordeel te onderbouwen (Goulden, 1992; Hunter et al., 1996). 'Primary Trait Scoring' bevindt zich op het midden van het continuüm en wordt door Hunter et al. (1996) onder de

Continuum of Scoring Methods



Opmerking. Overgenomen uit "The use of holistic versus analytic scoring for large-scale assessment of writing" door D. M. Hunter, M. Jones & B. S. Randhawa, 1996, The Canadian Journal of Program Evaluation, 11, p. 64. Copyright 1996.

Figuur 1. Continuüm met holistische en analytische benaderingen

holistische benaderingen geschaard. Vergelijkbaar met 'holistic scoring' worden de criteria voorafgaand bepaald, echter worden deze met meer precisie en exclusiviteit vastgesteld.

Aanleiding voor holistisch beoordelen bij Saxion Parttime School (SPS)

In de afgelopen jaren lag bij SPS de focus vooral op het beoordelen van afzonderlijke leeruitkomsten (Kader 2) met een voorgeschreven beroepsproduct. Daarbij gebruikt de student de eigen praktijksituatie om de leeruitkomst aan te tonen. Om de toetsing nog flexibeler te maken en de meerwaarde van opgeleverde producten voor de praktijk te verhogen, is SPS bij de deeltijdopleiding Social Work (SW) een pilot gestart met het beoordelen van drie samenhangende leeruitkomsten, die vallen onder dezelfde Eenheid van Leeruitkomsten (EvL) (zie Tabel 1). Studenten mogen zelf kiezen met welk(e) product(en) zij deze leeruitkomsten tegelijk aantonen in een portfolio met daaraan gekoppeld een assessment. SPS heeft vervolgens de beoordelingsmethodiek hierop aangepast en is deze samenhangende leeruitkomsten holistisch gaan beoordelen, volgens de variant 'holistic scoring' (Hunter et al., 1996) gebaseerd op de methode die gebruikt wordt bij het valideren van werkervaring (zie Kader 3). Daar vindt de beoordeling al holistisch plaats op basis van de (eenheden van) leeruitkomsten. Deze (eenheden van) leeruitkomsten vormden een mooie basis om de pilot vorm te geven. Hiermee wordt meer recht gedaan aan de kwaliteit van de prestatie zoals deze ook geleverd en beoordeeld wordt in de beroepspraktijk. Een student ziet meteen wat de kwaliteit van de gehele prestatie dient te zijn. Daarnaast wordt de creativiteit van de student nog meer gestimuleerd omdat de rubric niet meer als een afvinklijst ingezet kan worden om het product op te leveren.

Leeruitkomsten

SPS werkt met leeruitkomsten. Een leeruitkomst beschrijft wat een student geacht wordt te weten, te begrijpen en te kunnen toepassen na afronding van een leerperiode (NVAO, 2015). Leeruitkomsten worden onafhankelijk beschreven maar moeten voldoende concreet en specifiek zijn zodat er ruimte overblijft voor studenten om zelf te bepalen op welke manier ze de leeruitkomsten willen aantonen (Tuning project, z.d.). Toetsing kan daarmee ook leerwegaafhankelijk worden ingezet. Studenten en cursisten kunnen zonder aanwezigheidsverplichting en onafhankelijk van andere studenten en cursisten de toets voorbereiden en maken.

Het valideren van werkervaring bij SPS

Als een professional ruime werkervaring heeft en één of meerdere (eenheden van) leeruitkomsten kan aantonen met beroepsproducten kan een validerings-traject worden gestart. De professional verzamelt bewijsstukken die moeten voldoen aan de VRAAK-criteria (Variatie, Relevantie, Authenticiteit, Actualiteit, Kwantiteit). Tijdens de beoordeling van het portfolio bekijken twee assessoren afzonderlijk in hoeverre de professional de (eenheden van) leeruitkomsten heeft aangetoond. Op basis hiervan stellen ze gerichte vragen ter verduidelijking in een daarop volgend assessment. Op basis van een gemeenschappelijke beoordeling, gebaseerd op het portfolio en het assessment, komen de assessoren samen tot een eindoordeel of de (eenheden van) leeruitkomsten zijn aangetoond.

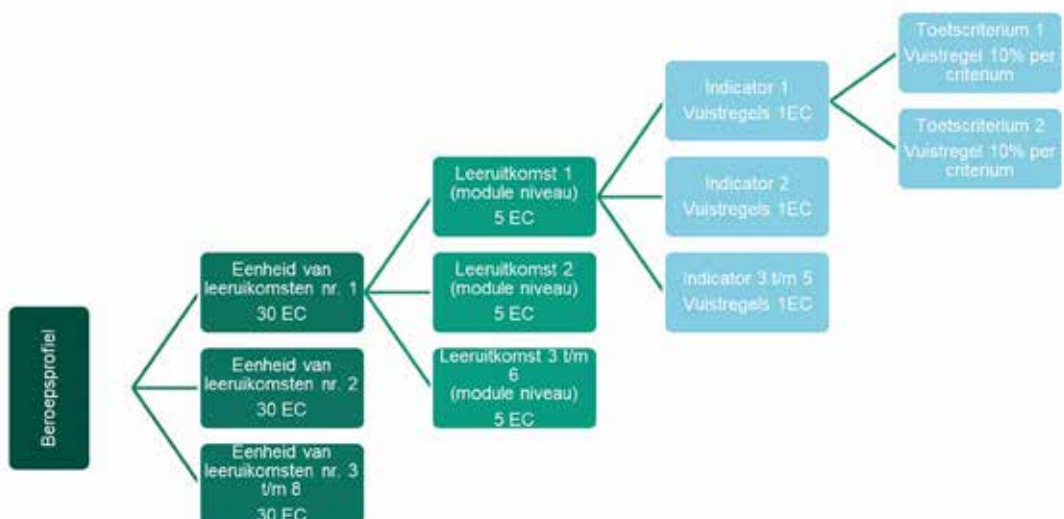
Waarom past holistisch beoordelen bij flexibel deeltijdonderwijs?

Een holistische beoordelingsschaal kan passend zijn bij producten waarin de prestatie moet worden getoond en waarbij de werkwijze of toepassingscontext per student verschilt (Bruijns & Kok, 2014). Nitko (2001) en Mertler (2000) benoemen dat een holistische beoordelingsschaal ook passend kan zijn bij opdrachten waar creativiteit in de oplossing is vereist en er meerdere oplossingsrichtingen zijn om een vraagstuk op te lossen. Holistisch beoordelen past daarom bij het onderwijsconcept van SPS, omdat studenten de samenhangende leeruitkomsten aantonen door middel van één of meerdere zelfgekozen beroepsproducten. Studenten kunnen deze producten zo vormgeven dat het optimaal aansluit bij de eigen organisatie. Centraal staat wat een student oplevert en niet hoe, in hoeveel tijd en waar een student dat leert. Er ligt een sterke focus op de output. Voordat wordt beschreven hoe het in een keer beoordelen van meerdere samenhangende leeruitkomsten er daadwerkelijk uit ziet, wordt eerst stilgestaan bij de beoordeling van een afzonderlijke leeruitkomst.

Het beoordelen van afzonderlijke leeruitkomsten

Bij SPS zijn de sector- en clustermodulen door veel verschillende opleidingen ontwikkeld en opgenomen in het curriculum. Daarom is het nog niet vrijlaten van de keuze van het product, maar ook de huidige beoordelingsmethodiek een bewuste keuze geweest om eerst een gemeenschappelijke beoordelingstaal te ontwikkelen.

In Figuur 2 is gevisualiseerd hoe bij SPS de vertaling van het landelijk opleidingsprofiel naar een dekkende set van leeruitkomsten eruit ziet. Elke module kent één leeruitkomst van 5 ECTS (Europees systeem voor de overdracht of accumulatie van studiepunten voor een opleiding of systeem) of een veelvoud daarvan. Deze leeruitkomst is vertaald naar een set van indicatoren (± 5) en deze zijn daarmee een concretisering van de leeruitkomst. Indicatoren beschrijven wat de student na afronding van de module kent en kan, en zijn een samenstelling van vaardigheden, zelfstandigheid, verantwoordelijkheid en benodigde kennis. De indicatoren zijn uitgewerkt in meerdere toetscriteria ($\pm$



Figuur 2. Van beroepsprofiel naar toetscriteria

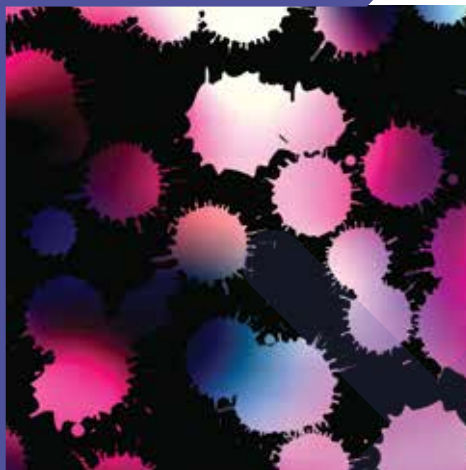
2 per indicator), waardoor een module (een leeruitkomst) gemiddeld wordt beoordeeld aan de hand van ± 10 toetscriteria. Elk toetscriterium betreft een specifieke kwaliteitseis die geldt voor het in de module voorgeschreven beroepsproduct. Bij de ontwikkeling zijn deze richtlijnen over het aantal indicatoren en toetscriteria bewust meegegeven om daarmee de focus bij de docenten al te leggen op de kernaspecten van de prestatie die door de student geleverd moet worden (zoals de leeruitkomst beoogt). Dit, om niet te verzanden in de formulering van een hele lijst met toetscriteria, zoals veel docenten vanuit hun eigen opleiding gewend waren.

De indicatoren en de toetscriteria zijn uitgewerkt in een rubric met een vijfpuntsschaal om meer te kunnen differentiëren tussen de verschillen in prestaties van studenten. Om tot het eindoordeel te komen worden alle afzonderlijke scores op de toetscriteria bij elkaar opgeteld en vindt er een eindoordeel plaats. De beoordeling zit daarmee aan de analytische kant (analytic scoring) van het continuüm (Hunter et al., 1996). Tabel 1 toont een voorbeeld van een uitwerking van een deel van de beoordeling van een module.

Het holistisch beoordelen van meerderesamenhangende leeruitkomsten

Een holistisch meetinstrument kan op verschillende manieren worden vormgegeven (Goulden, 1992). In deze pilot is ervoor gekozen om in een rubric de beoordelingsniveaus kort uit te schrijven met daaraan gekoppeld de punten (zie Tabel 2) omdat, met betrekking tot het laatste, het een eis vanuit Saxion is om te komen tot cijfers. Door de beoordelingsniveaus uit te schrijven, worden beoordelaars ook ondersteund bij de gezamenlijke normbepaling en een onderbouwing van de beoordeling (Hunter et al., 1996).

Bij het holistische beoordelingsinstrument zie je dat de indicatoren niet nogmaals uitgewerkt zijn in toetscriteria, zoals bij het beoordelen van een afzonderlijke module (leeruitkomst)



Holistisch beoordelen doet recht aan het beoordelen van authentieke situaties uit de praktijk

wel het geval is. De leeruitkomst wordt in een keer in zijn geheel beoordeeld. De indicatoren die behoren bij elke leeruitkomst kunnen worden gezien als de kenmerken of criteria waarop de beschrijving bij de beoordelingsniveaus is gebaseerd. Tabel 2 toont een voorbeeld van de uitwerking. De student kan per leeruitkomst 10 punten behalen en daarmee in totaal 30. Om een voldoende te halen moet de student op elk van de drie leeruitkomsten minimaal 6 punten behalen. Er is geen compensatie mogelijk. De assessoren zijn nauw betrokken geweest bij de ontwikkeling en implementatie van het meetinstrument. In werksessies is samen met onderwijskundigen van SPS het meetinstrument opgesteld. Vervolgens hebben de assessoren een speciale training gevolgd om te zorgen dat de criteria helder zijn, bias wordt verminderd en de betrouwbaarheid van de beoordeling wordt verhoogd (Fahim & Bijani,

Tabel 1. Voorbeeld van het beoordelen van een afzonderlijke module met bijbehorende leeruitkomst

Leeruitkomst				
De SW-professional past een cliëntversterkende methodiek eclectisch-integratief toe in een gesprek met een kwetsbare burger en geeft aanbevelingen voor cliëntversterkend werken in de praktijk.				
Indicatoren				
1. Je onderzoekt en beschrijft diverse cliëntversterkende methodieken en maakt hieruit een keuze op basis van de hulpvraag van de kwetsbare burger(s) welke methode jij gaat inzetten in jouw beroepspraktijk. 2. Je verantwoordt op welke manier je de gekozen methode inzet om de regie van de kwetsbare burgers(s) te versterken en aan de hulpvraag te voldoen en plant de uitvoering hiervan. 3. Je past de methode afhankelijk van de actuele gesprekssituatie met de kwetsbare burger (s) eclectisch- integratief toe. 4. Je evalueert de toegepaste methode met de kwetsbare burger(s) en met je praktijkbegeleider, zodat je hierop kunt reflecteren. 5. Je geeft jouw visie over de geschiktheid van de toegepaste methode voor de praktijk van de Social worker op basis van verkregen inzichten en geeft aanbevelingen voor cliëntversterkend werken voor de praktijkplek.				
Voorbeeld van een toetscriterium bijhorende bij indicator 1.				
1.1 De verantwoording van de keuze voor inzet van de clientversterkende methodiek in de praktijk is beschreven.				
Beoordelingsniveaus in de rubric van toetscriterium 1.1.				
Je verantwoordt je keuze niet of je hebt geen keuze gemaakt. 0 punten	Je hebt een keuze gemaakt maar de methodiek en de onderbouwing sluit niet aan bij de hulpvraag van de kwetsbare burger(s) in jouw beroepspraktijk.	De keuze voor een methodiek is verantwoord vanuit verschillende perspectieven (theorie, praktijk, ervaring) en deze keuze past bij de hulpvraag van de kwetsbare burger(s).	Daarnaast wordt in de verantwoording duidelijk welke afwegingen waarom zijn gemaakt om tot de definitieve keuze te komen.	Bovendien wordt duidelijk of de gekozen methode ook in andere beroepscontexten van de Social work-zorg toepasbaar is en is het een samenhangend geheel.
	3 punten	6 punten	8 punten	10 punten

Tabel 2. Voorbeeld van het holistisch beoordelen van drie samenhangende leeruitkomsten

Drie samenhangende leeruitkomsten				
• Leeruitkomst 1: De Social Work-professional formuleert een visie op kwetsbaarheid en geeft aanbevelingen gericht op het optimaliseren van de zorg door professionals als voor informele zorg. • Leeruitkomst 2: De SW-professional past een cliëntversterkende methodiek eclecticisch-integratief toe in een gesprek met een kwetsbare burger en geeft aanbevelingen voor cliëntversterkend werken in de praktijk • Leeruitkomst 3: De SW-professional ontwikkelt een interventie om de inclusie in de eigen werkpraktijk te bevorderen waarbij hij/zij zich positioneert als kwartiermaker om een bijdrage te leveren aan een inclusieve samenleving.				
Beoordelingsniveaus in de rubric van leeruitkomst 2				
Er wordt niet duidelijk welke cliënt-versterkende methode er is toegepast.	Er is een cliënt-versterkende methodiek toegepast echter is het niet duidelijk hoe er eclecticisch integratief gewerkt wordt of er worden geen aanbevelingen voor cliënt-versterkend werken in de praktijk beschreven.	Uit het gesprek blijkt hoe eclecticisch en integratief de cliënt-versterkende methode is toegepast. Uit de verantwoording wordt duidelijk welke afwegingen zijn gemaakt voor deze keuze tijdens het gesprek en dit is navolgbaar en onderbouwd. De aanbevelingen voor de toepassing in de praktijk zijn realistisch en passend.	Daarnaast blijkt uit de onderbouwing duidelijk de afwegingen die zijn gemaakt, waarbij concrete voorbeelden uit de praktijk zijn toegevoegd. Ook is aandacht besteed aan de implementatie in de eigen praktijk.	Tevens komt uit de onderbouwing naar voren hoe de keuze voor de toepassing van cliëntversterkend werken in een andere beroeps-context/praktijk van de Social worker-zorg kan worden toegepast.
0 punten	3 punten	6 punten	8 punten	10 punten

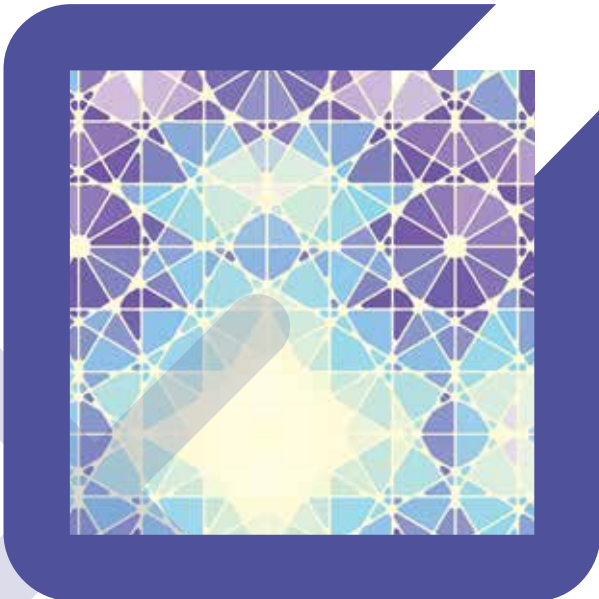
2011; Kang et al., 2019). Door middel van een aanvullende presentatie over het proces van aanmelding tot beoordeling zijn docenten vervolgens verder geïnstrueerd. Tijdens het proces waren onderwijskundigen beschikbaar voor vragen en monitoring van het proces. Zowel docenten als studenten vinden het traject en de wijze van beoordelen echt van meerwaarde. Het prikkelt om 'out of the box' te denken, omdat er geen lijst is met allerlei voorgeschreven criteria. De creativiteit wordt meer aangewakkerd en het product is nog beter afgestemd op de eigen praktijk.

Tot slot

In dit artikel is een voorbeeld gegeven hoe SPS naast het apart beoordelen van leeruitkomsten met meerdere indicatoren en toetscriteria het mogelijk heeft gemaakt om samenhangende leeruitkomsten in een of meerdere producten holistisch te beoordelen. Door deze leeruitkomsten holistisch te beoordelen is de verwachting dat studenten pro-actiever en creatiever worden in het vinden van een passende oplossing die nog beter afgestemd is op de eigen praktijk. Dit is iets wat zowel door studenten en docenten als belangrijk wordt ervaren. Een volgende stap is om het mogelijk te maken dat studenten ook voor de moduletoets een zelfgekozen product kunnen aanleveren. Daarnaast wordt gekeken hoe het holistisch beoordelen van meerdere samenhangende leeruitkomsten ook op grotere schaal voor meer opleidingen mogelijk gemaakt kan worden. Beide aspecten vragen, naast de expertise van docenten en assessoren, ook aanpassing binnen de logistieke systemen van hogescholen. Wanneer meerdere opleidingen deelnemen zal de opzet van zowel de beoordelingsmethodiek als het -proces steeds opnieuw geëvalueerd moeten worden, vooral wanneer er leeruitkomsten gecombineerd worden die in het curriculum van meerdere opleidingen zitten. Dit alles vraagt weer verdere aanvullende maatregelen om de kwaliteit te blijven borgen, zoals het aanscherpen of aanvullen van de training, het komen tot een gemeenschappelijke professionele norm, kalibratiesessies, et cetera.

Dankwoord

Ik wil de docenten van de deeltijdopleiding Social Work, Martijn Gunnewijk, Glenn Ripassa, Jacqueline Thuss, bedanken voor het meedenken, de bijdrage en prettige samenwerking vooraf, tijdens en na de pilot. ■



**Holistisch beoordelen
stimuleert nog meer de
creativiteit van studenten
om passende oplossingen
te vinden omdat de rubric
niet meer als een afvinklijst
ingezet kan worden om het
product op te leveren**

Referenties

- Barylá, E., Shelley, G., & Trainor, W. (2012). Transforming Rubrics Using Factor Analysis. *Practical Assessment, Research, and Evaluation*, 17, 4. <https://doi.org/10.7275/5nt2-g458>
- Cooper, C. R. (1977). Holistic evaluation of writing. In C. R. Cooper & L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging* (pp. 3-31). National Council of Teachers of English.
- Fahim, M., & Bijani, H. (2011). The effects of rater training on raters' severity and bias in second language writing assessment. *Iranian Journal of Language Testing*, 1(1), 1-16.
- Goulden, N. R. (1992). Theory and vocabulary for communication assessments. *Communication Education*, 41(3), 258-269. <https://doi.org/10.1080/03634529209378887>
- Hunter, D. M., Jones, R. M., & Randhawa, B. S. (1996). The use of holistic versus analytic scoring for large-scale assessment of writing. *The Canadian Journal of Program Evaluation*, 11(2), 61.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational research review*, 2(2), 130-144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Kang, O., Rubin, D., & Kermad, A. (2019). The effect of training and rater differences on oral proficiency assessment. *Language Testing*, 36(4), 481-504. <https://doi.org/10.1177/0265532219849522>
- Lloyd-Jones, R. (1977). Primary trait scoring. In C.R. Cooper & L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging* (pp. 33-68). National Council of Teachers of English.
- Mertler, C. A. (2000). Designing scoring rubrics for your classroom. *Practical assessment, research, and evaluation*, 7(1), 25. <https://doi.org/10.7275/gcy8-0w24>
- Nitko, A. J. (2001). *Educational assessment of students* (3rd ed.). Merrill.
- Sadler, D. R. (2009). Transforming holistic assessment and grading into a vehicle for complex learning. In G. Joughin (Ed.), *Assessment, learning, and judgement in higher education* (pp. 45-64). Springer.
- Schunn, C., Godley, A., & DeMartino, S. (2016). The reliability and validity of peer review of writing in high school AP English classes. *Journal of Adolescent & Adult Literacy*, 60(1), 13-23. <https://doi.org/10.1002/jaal.525>
- Tuning Project. (z.d.). *Reference Points for the Design and Delivery of Degree Programmes in European Studies*. https://www.unideusto.org/tuningeu/images/stories/Publications/EUROPEAN_STUDIES_FOR_WEBSITE.pdf