

Rubrics

Albert van der Kaap, 2016

Really, the greatest rubric in the world is worth little if there isn't a conscientious teacher consistently drawing the students' attention to it and making achievement of the goals attainable for the students who are willing to work at it (Paul Jolly).

The best feedback is worthless if students are not coached to be evaluators. They must learn how to evaluate the feedback that the teacher provides, while also self-assessing (Mark Barnes).

1. Inleiding

A rubric is a scaled set of criteria that clearly defines for the student and the teacher what a range of acceptable and unacceptable performances looks like. Its purpose is to provide a description of successful performances. A critical feature of rubrics is language that describes rather than labels performance. Evaluative words, like 'better', 'more often', and 'excellent' do not appear in rubrics. Instead, the language must precisely define actions in terms of what the student actually does to demonstrate skill or proficiency at that level.

AASL and AECT, 1998 p. 177

Een rubric is een instrument om verschillende soorten prestaties van leerlingen te evalueren of te beoordelen. Je kunt dan denken aan het doen van onderzoek, het maken van een ontwerp, het houden van een presentatie en ook het proces van samenwerken. Zoals uit bovenstaande definitie duidelijk wordt moet een rubric, om een echte rubriek genoemd te mogen worden, aan bepaalde eisen voldoen. De criteria moeten in eenduidige taal op verschillende niveaus (meestal vier) worden beschreven. Woorden als 'beter', 'vaak', of 'uitstekend', of 'mager' moeten vermeden worden. In plaats daarvan moeten de omschrijvingen nauwkeurig het verlangde gedrag van leerlingen beschrijven.

Rubrics kunnen zowel door docenten als door leerlingen gebruikt worden.

Rubrics voor docenten

Een rubric wordt vaak gebruikt voor het beoordelen van prestaties van leerlingen omdat het houvast biedt bij het beoordelen van het proces en het product (inzicht geeft in het beheersingsniveau m.b.t. tot de vereiste vaardigheden en/of competenties). Hierdoor wordt de objectiviteit van de beoordeling vergroot (Moskal, 2000). Voorwaarde is dan wel dat de criteria zodanig zijn geformuleerd dat ze zich niet lenen voor persoonlijke interpretatie van de beoordelaars.

Een rubric kan ook aan de voorkant van het leerproces goede diensten bewijzen. Door de eisen die aan een opdracht worden gesteld nauwkeurig van tevoren vast te leggen in leerdoelen en in de bijbehorende rubric, kan deze helpen om een zo helder mogelijke

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

instructie voor de opdracht te schrijven. De geformuleerde criteria zorgen ervoor dat duidelijk wordt:

- wat er precies van leerlingen verwacht wordt tijdens het werken aan de opdracht;
- aan welke eisen het uiteindelijke product (met betrekking tot vaardigheden/competenties) moet voldoen.

Rubrics voor leerlingen

Aanleren

Rubrics kunnen ook het leren van leerlingen ondersteunen. We spreken dan van instructional rubrics. Rubrics waarin aangegeven wordt op welke aspecten en volgens welke criteria een prestatie beoordeeld zal worden, maken leerlingen niet alleen duidelijk wat er van hen verwacht wordt, maar ook hoe zij zich kunnen verbeteren. Dat dát zeer effectief kan zijn blijkt uit de volgende opmerking van een leerling: "Als je iets niet goed hebt gedaan kan de docent aantonen dat je had kunnen weten wat er van je verlangd werd" (Marcus, 1995).

Dit geeft aan dat het belangrijk is leerlingen vanaf het begin te betrekken bij het beoordelingsproces. Door met hen vooraf de criteria te bespreken die bij de beoordeling van belang zijn, zijn zij beter op hun taak voorbereid en zullen zij beter in staat zijn de opdrachten zelfstandig uit te voeren (Rochford en Borchert, 2011). Het ligt daarom voor de hand om de rubric op te nemen in de instructie voor de leerlingen (Wiggins, 1998).

Als rubrics in een voor de leerlingen begrijpelijke taal beschrijven waaraan hun prestatie moet voldoen, kunnen ze leerlingen helpen het resultaat van hun activiteiten te evalueren (formatieve evaluatie). Mocht daar aanleiding toe zijn, dan kunnen ze maatregelen treffen om de, nog ontoereikende, prestatie bij te stellen om die alsnog aan de vereisten te laten voldoen. Leerlingen krijgen inzicht in de kwaliteit van hun eigen werk en dus ook in hun eigen leerproces.

Leerlingen kunnen een rubric ook summatief gebruiken. Aan de hand van de rubric kunnen zij zowel zichzelf (self-assessment) als hun medeleerlingen (peer-assessment) beoordelen.

Het Algemeen Pedagogisch Studiecentrum (APS) heeft een aantal jaren geleden onderzoek gedaan naar het gebruik van rubrics. Het instituut vat de onderzoeksresultaten in *'Het mes snijdt aan zes kanten'* als volgt samen:¹

- Een rubric leidt tot een objectievere beoordeling van leerlingenwerk.
- Een rubric biedt houvast bij het nakijken van een product of bij het observeren van het uitvoeren van de vaardigheid.
- Een rubric geeft de mogelijkheid een beoordeling duidelijk naar leerlingen (en ouders) te verantwoorden.
- Meer dan de helft van de leerlingen denkt dat hij of zij een kwalitatief beter werkstuk heeft kunnen produceren door het gebruik van een rubric.
- Alle leerlingen vinden dat door gebruik te maken van een rubric zij hun tijd effectiever hebben besteed.
- Leerlingen hechten veel waarde aan de houding van de docent ten opzichte van een gegeven opdracht. Wanneer je als docent aan de leerlingen laat zien dat je de

¹ <http://www.excellentrekenen.nl/media/1673/rubrics.pdf>

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>



opdracht, de wijze waarop deze uitgewerkt wordt en de kwaliteit van de beoordeling belangrijk vindt, gaan de leerlingen hier ook serieus mee om.

- Leerlingen vinden het prettig als gericht werken tot een hoog cijfer leidt. Dat kan met een rubric.
- Het is niet gebleken dat het nakijken met rubrics sneller gaat. Wel is gebleken dat het nakijken zeker niet langer duurt. De tijd die het kost om een praktische opdracht na te kijken is afhankelijk van de opdracht en van de docent zelf.
- Het ontwikkelen van een rubric kost, zeker de eerste keer, veel tijd, die je echter later vaak ook aan tijd terugwint. Bovendien zie je de inspanningen vaak terug in de kwaliteit van het werk van de leerlingen.
- Niet vakspecifieke rubrics zijn - al of niet aangepast - vakoverstijgend toepasbaar, wat de transfer van vaardigheidsonderwijs van het ene naar het andere vak bevordert.

2. Hoe werkt een (analytische) rubric?

Een rubric kenmerkt zich door:

- Een overzicht van criteria waaraan een te leveren prestatie wordt afgemeten. Deze worden afgeleid van de leerdoelen.
- De aanduiding van een aantal niveaus van beheersing. Meestal gaat het om vier niveaus met kwalificaties als 'van zeer zwak tot uitstekend', 'van beginner tot expert'. Soms kiest men, om pedagogische redenen, ook voor de omgekeerde volgorde.
- Een (gewoonlijk zeer beknopte) beschrijving van de onderscheiden beheersingsniveaus. Bij een 'zuivere' rubric zijn deze descriptoren verplicht en essentieel (zie de definitie aan het begin van dit hoofdstuk). Zij geven nauwkeurige informatie over de verwachte prestatie.

Een rubric heeft meestal een vaste vorm:

- Een kolom '**Hoofdrubriek**' die aangeeft op welke vaardigheid/competentie de rubric betrekking heeft. Bijvoorbeeld: samenwerking, communicatie, verslaglegging of presenteren.
- Een kolom '**Subrubriek**'. De hoofdrubriek kan, desgewenst, in een aantal subrubrieken worden onderverdeeld.
- Een kolom '**Criteria**' waarin per (sub)rubriek wordt vermeld waarop het werk wordt beoordeeld.
- Een kolom waarin op elk criterium **gescoord** kan worden.
- Een kolom waarin, indien gewenst, wordt vermeld hoe zwaar elk criterium **gewogen** wordt, dat wil zeggen meetelt in de uiteindelijke beoordeling.

Bij het maken van een rubric is het zaak om onduidelijke kwalificaties, onheldere taal en negatief taalgebruik te vermijden.

In schema ziet een rubric er als volgt uit:

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Hoofdrubriek	Subrubriek	Criteria	1	2	3	4	Score	Weging
1								
3								
4								
5	1							
	2							
	3							
	4							
	5	1						
		2						
		3						
		4						
		5						

Afhankelijk van de uitgebreidheid van de rubric en het doel waarvoor hij gebruikt wordt, zie je in de praktijk vaak dat de eerste twee en de laatste twee kolommen ontbreken.

De consistentie in de beoordeling met rubrics kan worden versterkt door voorbeelden te geven van elk bedoeld niveau, bij voorkeur ontleend aan werk van de leerlingen. Dat zou kunnen in de vorm van een zogeheten 'anker'. Een anker is een item dat in twee of meer prestaties van leerlingen voorkomt en dat als ijkpunt wordt genomen om verschillen in moeilijkheidsgraad/uitvoeringsniveau te kunnen constateren. Het is een middel om prestaties vergelijkbaar te maken. Dergelijke ankers helpen niveaubeschrijvingen te concretiseren en fungeren daarmee als een soort ijkpunt voor de beoordelaar.

Hieronder staat een rubric die is ontwikkeld voor het onderdeel verspringen van het vak Bewegen, Sport en Maatschappij (BSM). Terwijl de grens onvoldoende/voldoende bij veel rubrics bij de overgang van niveau twee naar niveau drie ligt, is hier uit een oogpunt van motivatie, bewust gekozen voor slechts één onvoldoende niveau. Overigens worden voor elk niveau aanvullend filmpjes gemaakt om de verschillende niveaus te visualiseren.

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

	Onvoldoende (tot 5,5)	Voldoende (5,5 - 7,0)	Goed (7,0 - 8,5)	Litstekend (8,5 - 10)
1. De aanloop	Komt verkeerd uit bij de afzetbalk (te ver af of ongelidig over de afzetbalk); versnelt niet of nauwelijks.	Komt met afgepaste passen in de buurt van de afzetbalk (soms er overheen); versnelt in lichte mate; ontwikkelt een redelijke snelheid.	Komt met afgepaste passen precies uit voor de afzetbalk; loopt versnellend aan; ontwikkelt een hoge snelheid.	Komt met afgepaste passen precies uit voor de afzetbalk; loopt versnellend aan; ontwikkelt een hoge loopsnelheid; maakt de een na laatste pas iets groter voor een betere verticale afzet.
2. De afzet	Bij de afzet blijft het afzetbeen lager dan heuphoogte.	Komt bij de afzet licht omhoog; zwaait het zwaaibeen omhoog tot heuphoogte.	Komt bij de afzet duidelijk omhoog; zwaait actief het zwaaibeen omhoog, zeker tot heuphoogte.	Strekt enkel en knie van het afzetbeen krachtig; komt bij de afzet krachtig omhoog; geeft de afzet een extra impuls met het actief en krachtig opzwaaien van het zwaaibeen.
3. De zweeffase	Komt in de zweeffase nauwelijks tot een hurkhouding.	Maakt in de lucht een hurkbeweging op het hoogste punt; zwaait de onderbenen actief voorwaarts voor de landing.	Maakt in de eerste opgaande zweeffase een hurkbeweging; zwaait de benen vlak voor de landing krachtig voorwaarts.	Maakt in de lucht een loop(sprong) beweging (hitch-kick) ter ondersteuning van de afzet; zwaait de benen actief voorwaarts na het hoogste punt.
4. De landing	Zet de voeten snel in het zand; gaat bij de landing achter de voeten zitten.	Landt met de voeten voor het zwaartepunt; hurkt dicht bij de voeten.	Landt met de voeten voor het zwaartepunt; hurkt bij de landing snel in of draait zijwaarts om een grotere afstand te overbruggen.	Grijpt met de voeten ver voorwaarts in het zand; hurkt bij de landing snel in of draait zijwaarts weg om een grotere afstand te overbruggen.
5. Het coachen	Is vooral met zijn/haar eigen bewegingsverloop bezig; is moeilijk te coachen.	Kan met gerichte instructie analytisch naar het bewegingsverloop van een medeleerling kijken en daar een aanwijzing op geven.	Kan met gerichte instructie meer analytisch naar het bewegingsverloop van een medeleerling kijken en daar een aanwijzing op geven.	Kan analytisch naar het bewegingsverloop van een medeleerling kijken en daar een passende aanwijzing en een passend voorbeeld bij geven.

Wel of geen beschrijving van de verschillende niveaus?

Heel vaak is het moeilijk om elk van de vier beheersingsniveaus op een betekenisvolle manier te beschrijven. Men neemt al gauw zijn toevlucht tot beschrijvingen als slecht, matig, redelijk en goed of nooit, soms, af en toe en vaak. In dat geval kan men ervoor kiezen om alleen het laagste en het hoogste beheersingsniveau te beschrijven en het werk van de leerlingen op elk criterium te scoren op een schaal van 1-4 (of eventueel 1-6 of 1-8 als een dergelijke nuancering mogelijk en zinvol is). Meestal wordt bewust gekozen voor een even aantal niveaus om de beoordelaar te dwingen een duidelijke keuze te maken tussen voldoende en onvoldoende. Dit veronderstelt echter wel dat de beoordelaar in staat is onderscheid te maken tussen niveau 2 en 3. Als dat niet mogelijk is, kan ook voor een driepuntsschaal worden gekozen. In dat geval is wel of niet aan het criterium voldaan of is de prestatie op dit criterium twijfelachtig.

Hoofdrubriek	Subrubriek	Criteria	Beschrijving	1	2	3	4	Beschrijving	Score	Weging
1										
3										
4										
5	1									
	2									
	3									
	4									
	5	1		O	O	O	O			
		2		O	O	O	O			
		3		O	O	O	O			
		4		O	O	O	O			
		5		O	O	O	O			

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Een stap verder gaat het model waarbij het criterium alleen op het hoogste niveau wordt omschreven. Een dergelijke schaal wordt meestal gebruikt in enquêtes en opiniepeilingen waarbij de geënquêteerden moeten reageren op een stelling. Een dergelijk schaal is echter ook goed bruikbaar voor het beoordelen van vaardigheden op criteria die moeilijk in concreet waarneembaar gedrag te beschrijven zijn. Bij een dergelijke schaal geeft de beoordelaar aan in hoeverre hij het eens is met een uitspraak met betrekking tot een bepaald criterium.

- 1= sterk mee oneens
- 2= mee oneens
- 3= mee eens
- 4= sterk mee eens

Voor het beoordelen van vaardigheden ligt een beoordeling van slecht tot goed meer voor de hand.

- 4= goed/volledig
- 3= voldoende
- 2= onvoldoende
- 1= zeer onvoldoende

Criteria	Beschrijving	1	2	3	4
1		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Single-point rubric

In plaats van een meerpuntsschaal, kan ook gekozen worden voor een eenpuntsschaal (single-point rubric), waarbij het niveau, zoals traditioneel in de derde kolom (van een vierpuntsschaal) wordt beschreven, het uitgangspunt is. Deze rubric is minder geschikt voor een summatieve beoordeling, waarbij een gedetailleerde beoordeling, vaak met een cijfer met één decimaal wordt gevraagd.

Voor een formatieve beoordeling, gericht op het verbeteren van de prestaties van leerlingen, biedt een eenpuntsubric de mogelijkheid om, uitgaande van een voldoende prestatie, aan te geven op welke criteria de leerling nog tekort schiet en wat hij moet doen om zich te verbeteren. Ook kan per criterium, waar relevant, precies worden aangegeven waarom de prestatie excellent is, of in ieder geval beter dan gemiddeld.

Hoewel het invullen van deze rubric tamelijk tijdrovend kan zijn, geeft hij de leerlingen veel inzicht in vooral de zwakke onderdelen van zijn prestatie, zodat hij weet hoe hij zich juist op die punten kan verbeteren.

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Waarin schiet de prestatie tekort?	Criterium voor een voldoende prestatie	Waarin stijgt de prestatie uit boven het vereiste niveau?
	Criterium 1	
	Criterium 2	
	Criterium 3	
	Criterium 4	

4. Holistische rubric

Een laatste vorm van een rubric is de zogenaamde holistische rubric. In deze rubric wordt, in algemene termen, een aantal niveaus van presteren beschreven. Per niveau moet worden bepaald wanneer dit niveau is bereikt. De niveaus kunnen worden beschreven met cijfers, letters of in woorden (van beginner tot gevorderd).

Breakfast in Bed: Holistic Rubric

Score	Description
4	All food is perfectly cooked, presentation surpasses expectations, and recipient is kept exceptionally comfortable throughout the meal.
3	Food is cooked correctly, the meal is presented in a clean and well-organized manner, and the recipient is kept comfortable throughout the meal.
2	Some food is cooked poorly, some aspects of presentation are sloppy or unclean, or the recipient is uncomfortable at times.
1	Most of the food is cooked poorly, the presentation is sloppy or unclean, and the recipient is uncomfortable most of the time.

De holistische rubric heeft veel weg van de **Structure of the Observed Learning Outcomes (SOLO)** taxonomie van Biggs en Collis (1982). De SOLO-taxonomie onderscheidt vijf niveaus van presteren:

- Pre-structureel.** De taak wordt onjuist benaderd en men heeft de opdracht helemaal niet begrepen.
- Unistruktureel.** Slechts één aspect uit een meervoudige taak wordt opgepakt en uitgewerkt.
- Multi-structureel.** Meerdere aspecten uit de taak worden los van elkaar benoemd en opgesomd, zonder enige relatie daartussen weer te geven.
- Relationeel.** De componenten uit de taak worden op een samenhangende manier aangepakt en de betekenis van het geheel wordt geduid.
- Uitgebreid abstract.** Het geïntegreerde geheel van het relationele niveau wordt geabstraheerd en toegepast in andere contexten (transfer).

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Dit beoordelingsinstrument is vooral geschikt voor (grotere) opdrachten waarin sprake is van complexe relaties en/of waarin onderscheid gemaakt kan worden tussen concrete en meer abstracte antwoorden.

Een voorbeeld

Bron

Uit een boek van een Sovjethistoricus in 1968:

„Het Marshallplan van 1947 leidde tot een teruggang in de handel van de SU. De VS hoopten dat het plan zou leiden tot verdeeldheid onder de communistische staten en dat deze staten onder Amerikaanse invloed zouden komen. Het was ook duidelijk dat veel van het Marshallplan bedoeld was om de militaire macht van West-Duitsland te doen herleven.”²

Gebruik de bron

Hoe waardevol is deze bron om meer te weten te komen over het Marshallplan? (Hoe waardevol is deze bron voor onderzoek naar het Marshallplan?)

Normeringsmodel volgens de SOLO-taxonomie.

Niveau	Omschrijving
1. Prestructureel	De leerling noemt een enkel element, maar geeft er geen blijk van begrijpen wat er van hem wordt verwacht.
2. Unistruktureel	De leerling gaat alleen in op elementen uit de bron zelf, maar koppelt deze niet aan relevantie, betrouwbaarheid en representativiteit.
3. Multistruktureel	De leerling gaat in op elementen uit de bron zelf en koppelt deze aan relevantie, betrouwbaarheid en representativiteit.
4. Relationeel	De leerling gaat in op elementen uit de bron zelf, koppelt deze aan relevantie, betrouwbaarheid en representativiteit en geeft er blijk van oog te hebben voor relevante informatie die de bron niet bevat.
5. Extended abstract	De leerling gaat in op elementen uit de bron zelf, koppelt deze aan relevantie, betrouwbaarheid en representativiteit en geeft er blijk van oog te hebben voor relevante informatie die de bron niet bevat. Hij plaatst zijn redenering in het bredere kader van de Koude Oorlog.

² Deze opdracht is afkomstig van

<http://www.geschiedenisindeklas.com/download.php?file=cGRmL2xlc21hdGVyaWFhC9vZWYyYW50LnZyYWcuLnBkZg>

Een beoordelingsmodel dat hierop lijkt is het 'banded-marking' of 'levels-based' model. Dit model is vooral geschikt voor vragen waarvoor leerlingen veel punten (meer dan zes) kunnen behalen en waarop verschillende antwoorden mogelijk zijn. Voor dergelijke vragen is dit een betrouwbaarder beoordelingsmodel dan een gedetailleerd beoordelingsmodel (Bramley, 2008).

Normeringsmodel van dezelfde opdracht volgens het 'banded-marking' of 'levels-based' model.

<p>9-11 punten De leerling beschrijft alle informatie die de bron geeft over het Marshallplan. Hij vermeldt bovendien voor welke vraag deze informatie waardevol is. Ook gaat hij in op de vraag wat de bron minder waardevol/betrouwbaar maakt. Hij gaat daarnaast in op aspecten van het Marshallplan waarover de bron geen informatie geeft. Hij vermeldt dat de waarde van een bron wordt bepaald door de vraag die je hebt over het Marshallplan.</p>
<p>5-8 punten De leerling beschrijft alle informatie die de bron geeft over het Marshallplan. Hij vermeldt bovendien voor welke vraag deze informatie waardevol is. Ook gaat hij in op de vraag wat de bron minder waardevol/betrouwbaar maakt.</p>
<p>1-4 punten De leerling beschrijft de informatie die de bron geeft over het Marshallplan. Hij gaat niet in op de vraag in hoeverre de bron waardevol is.</p>
<p>0 punten De leerling heeft de vraag niet beantwoord.</p>

In sommige gevallen kan dit model echter ook gebruikt worden voor een meer gedetailleerde beoordeling. Elementen die in de uitvoering van de opdracht aan de orde moeten zijn gesteld staan in de rubric vermeld in een oplopende moeilijkheidsgraad (van niveau 1 tot en met niveau 3). Elk van de drie niveaus is onderverdeeld in verschillende elementen. Per element is een punt te verdienen. De scores op de elementen tezamen weerspiegelen de mate waarin een leerling aan de gestelde doelen heeft voldaan. Dit model is met name ook geschikt voor een formatieve beoordeling. De leerling krijgt inzicht in het niveau dat hij aankan en krijgt zicht op die aspecten waarop hij zich kan verbeteren.

Meer gedetailleerd normeringsmodel van dezelfde opdracht volgens het 'banded-marking' of 'levels-based' model.

Beoordelingsformulier				
		Descriptor		Score
Niveau 1	a	Het Marshallplan leidde tot een teruggang in de handel van de S.U.	1	1-4
	b	De bedoeling van de V.S. was om verdeeldheid te stichten onder de communistische staten.	1	
	c	De bedoeling van de V.S. was om de communistische staten onder Amerikaanse invloed te brengen.	1	
	d	De bedoeling van de V.S. was om de militaire macht van West-Duitsland te doen herleven.	1	

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Niveau 2	a	De bron is waardevol omdat hij informatie geeft over de visie van de S.U. op het Marshallplan.	1	5-8
	b	De bron is minder waardevol/eenzijdig omdat deze alleen informatie geeft vanuit Russisch perspectief.	1	
	c	De bron is mogelijk niet representatief voor het Russische standpunt omdat het slechts gaat om de visie van één Sovjethistoricus op het Marshallplan.	1	
	d	De bron is minder waardevol/betrouwbaar omdat hij stamt uit 1968.	1	
Niveau 3	a	De bron is minder waardevol omdat hij geen informatie geeft over het Marshallplan vanuit Amerikaans perspectief.	1	9-11
	b	De bron is minder waardevol omdat hij geen informatie geeft over de inhoud van het Marshallplan.	1	
	c	Hoe waardevol de bron is hangt af van de historische vraag die je met betrekking tot het Marshallplan wilt beantwoorden.	1	

5. Rubric voor het beoordelen van rubrics

De ene rubric is de andere niet. Om tot een rubric te komen die aan de individuele eisen van de gebruiker beantwoordt en de toets der kritiek kan doorstaan, kan de bijgestelde versie van de rubric van Arter & Chappuis (Van Strien & Joosten-ten Brinke, 2016) voor het beoordelen van rubrics wellicht waardevolle diensten bewijzen.

	Goed (2 pnt)	Voldoende (1 pnt)	Onvoldoende (0 pnt)
Validiteit			
Dekkingsgraad	De inhoud weerspiegelt correct wat er van studenten realistisch gezien mag worden verwacht gezien hun niveau en de leerdoelen. De inhoud van de rubric geeft de best mogelijke weergave van wat er nodig is om een goede prestatie te leveren op de vaardigheid of het product dat wordt beoordeeld.	De inhoud weerspiegelt wat er van studenten realistisch gezien mag worden verwacht gezien hun niveau en de leerdoelen, al geldt dit niet voor alle criteria. De inhoud van de rubric geeft grotendeels een goede weergave van wat er nodig is om een goede prestatie te leveren, maar er zijn punten waarop verbetering nodig is.	De inhoud weerspiegelt niet wat er van studenten mag worden verwacht gezien hun niveau en de leerdoelen. De inhoud van de rubric geeft geen goede weergave van wat er nodig is om een goede prestatie te leveren, waardoor niet duidelijk is wat de rubric beoogt te meten.

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Prestatieniveau	Studenten mogen verondersteld worden in staat te zijn het hoogste niveau te bereiken op alle criteria. Om elk criterium op 'voldoende' niveau te beheersen moeten studenten een acceptabele prestatie laten zien.	Studenten mogen verondersteld worden in staat te zijn het hoogste niveau te bereiken op alle criteria. Op enkele criteria kan met een zwakke prestatie reeds een voldoende of hoger worden behaald.	De eisen die worden gesteld om het hoogste niveau te bereiken zijn onrealistisch en/of studenten kunnen door een zwakke prestatie op meerdere criteria reeds een voldoende of hoger halen.
Kwaliteit van de criteria			
Aantal, samenhang, opbouw	De rubric omvat precies het juiste aantal criteria, zodat de complexiteit van de te leveren prestatie wordt weerspiegeld. De criteria kennen een goede samenhang en logische opbouw.	Het aantal criteria vereist enige aanpassing, zoals het opsplitsen van een afzonderlijk criterium in twee aparte criteria, of het samenvoegen van twee criteria. De criteria kennen verder een goede samenhang, maar de opbouw is niet altijd logisch.	De rubric bestaat uit een lange lijst criteria die niet of nauwelijks met elkaar lijken samen te hangen en waarin geen keuzes lijken te zijn gemaakt. De opbouw is niet logisch.
Weging	Het is duidelijk bedoeld is voor formatieve en/of summatieve doeleinden. Indien summatieve doeleinden, dan is in de rubric expliciet aangegeven hoe zwaar elk criterium meetelt, waardoor is af te leiden welke criteria belangrijker zijn dan andere (eventueel in begeleidende tekst). Alle criteria krijgen de nadruk die zij verdienen.	Het is duidelijk of de rubric bedoeld is voor formatieve en/of summatieve doeleinden. Indien summatieve doeleinden, dan is in de rubric weliswaar expliciet (eventueel in begeleidende tekst) aangegeven hoe zwaar elk criterium meetelt en welke criteria belangrijker zijn dan andere, maar sommige criteria krijgen te veel of te weinig nadruk.	Het is niet duidelijk of de rubric bedoeld is voor formatieve en/of summatieve doeleinden. In de rubric wordt niet expliciet (ook niet in begeleidende tekst) aangegeven hoe zwaar elk criterium meetelt en welke criteria belangrijker zijn dan andere. De criteria krijgen hierdoor niet de nadruk die zij verdienen.

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Onafhankelijkheid	De criteria zijn onafhankelijk van elkaar, waardoor zij verschillende aspecten meten. Er is geen overlap tussen verschillende indicatoren. Wat hoort bij de ene indicator komt nergens anders terug.	De criteria zijn grotendeels onafhankelijk. Er is enige overlap tussen verschillende indicatoren, maar de rubric is desondanks in staat om verschillende aspecten te meten.	De criteria zijn niet onafhankelijk. Er is veel overlap tussen verschillende indicatoren. Verschillende indicatoren omvatten hetzelfde, waardoor de criteria niet verschillende aspecten meten.
Onderscheidend vermogen	Het aantal beoordelingsniveaus per criterium is logisch en toereikend. Er zijn voldoende niveaus om adequaat onderscheid te kunnen maken tussen goede en minder goede studenten en om voortgang te kunnen meten.	Het aantal beoordelingsniveaus per criterium is niet overal toereikend om adequaat onderscheid te kunnen maken tussen goede en minder goede studenten en om voortgang te kunnen meten, maar het aantal is redelijk eenvoudig aan te passen door één niveau toe te voegen dan wel twee niveaus samen te voegen.	Het aantal beoordelingsniveaus is veel te groot of juist te klein om zinvol en betrouwbaar onderscheid te kunnen maken tussen studenten. Grote aanpassingen zijn nodig.
Kwaliteit van de indicatoren			
Aanwezig	Elk beoordelingsniveau is voorzien van indicatoren.	Niet alle beoordelingsniveaus zijn voorzien van indicatoren.	Geen van de beoordelingsniveaus is voorzien van indicatoren.

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Transparantie en paralleliteit	De indicatoren zijn gedetailleerd en concreet genoeg zodat duidelijk is wat de student moet laten zien om een goede prestatie te leveren (ze zijn kwalitatief i.p.v. kwantitatief) en de beoordelaar een juiste beoordeling kan geven. Er is een logische opbouw van indicatoren over de niveaus (paralleliteit).	Er is een poging ondernomen om indicatoren gedetailleerd en concreet te formuleren, maar sommige bevatten nog enkele vage beschrijvingen. Ze zijn wel grotendeels kwalitatief (en niet kwantitatief). Er is veelal sprake van paralleliteit, maar bij een aantal criteria is de opbouw niet logisch.	De indicatoren zijn vaag en weinig concreet beschreven, waardoor niet duidelijk wordt wat er van de student wordt verwacht. Of er wordt alleen onderscheid gemaakt tussen beoordelingsniveaus met woorden als 'zeer', 'erg' en 'enige', of ze zijn volledig kwantitatief. De paralleliteit ontbreekt meestal.
---------------------------------------	---	--	---



Referenties

American Association of School Librarians and Association for Educational Communications and Technology. 1988. Information power: Guidelines for school library media programs. Chicago: American Library Association.

Arter, J.A., & Chappuis, J. (2006). *Creating & recognizing Quality Rubrics*. Boston: Pearson.

Biggs, J. B., & Collis, K., F. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome)*. New York, New York: Academic press, Inc.245.

Bramley, T. (2008). Mark scheme features associated with different levels of marker agreement. In *British Educational Research Association (BERA) Annual Conference*. Heriot-Watt University, Edinburgh, UK. Geraadpleegd van https://cerp.aqa.org.uk/sites/default/files/pdf_upload/CERP_RP_APM_24112011.pdf

González, J. (z.d.)Your Rubric Is a Hot Mess; Here's How to Fix It. Geraadpleegd van <http://www.brilliant-insane.com/2014/10/single-point-rubric.html>

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144.

Moskal, B.M. (2000). Scoring rubrics: What, when and how? *Practical Assessment and Evaluation*, 7(3). Verkregen op 20-09-2011 van <http://pareonline.net/getvn.asp?v=7&n=3>

Moskal, B.M., & Lydens, J.A. (2000) Scoring rubric development: Validity and reliability. *Practical Assessment and Evaluation*, 7(10). Verkregen op 20-09-2011 van <http://pareonline.net/getvn.asp?v=7&n=10>

Pinot de Moira, A. (2014). *Levels-based marking schemes and marking bias*. Geraadpleegd van https://cerp.aqa.org.uk/sites/default/files/pdf_upload/CERP_RP_APM_24112011.pdf

Rochford, L. en Borchert, P.S. (2011). Assessing Higher Level Learning: Developing Rubrics for Case Analysis in *Journal of Education for business*, 86: 258-265.

Stellmack, M.A., Konheim-Kalkstein, Y.L., Manor, J.E., Massey, A.R., & Schmitz, J.A.P. (2009). An assessment of reliability and validity of a rubric for grading APA-style introductions. *Teaching of Psychology*, 36, 102-107.

Strien, J. van, & Joosten-ten Brinke, D. (2016). Het beoordelen van de kwaliteit van rubrics. *Examens*, februari 2016, nr. 1. Geraadpleegd van https://www.ou.nl/Docs/Faculteiten/Welten/nieuws%202016/Van_Strien_Joosten-ten_Brinke_2016_het_beoordelen_van_de_kwaliteit_van_rubrics.pdf

Wiggins, Grant, en McTighe, Jay (1998). *Understanding by Design*. Expanded 2nd Edition. Alexandria, VA: ASCD, 2005.

Bron: SLO, Albert van der Kaap, <http://onderzoekinzestappen.slo.nl>

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.

Wiggins, G. (2012) Seven Keys to Effective Feedback. *In Educational Leadership*, September 2012 | Volume 70 | Number 1.

