

Tabel 1. De rubric voor het beoordelen van rubrics (bijgestelde versie van Arter & Chappuis, 2006)

	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.
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 of de rubric 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.
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.
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 ipv 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.

Elk criterium is onderverdeeld in subcriteria en alle criteria tellen even zwaar mee.