

**Competenties van vmbo-leerlingen beoordelen met behulp van
rubrics: mogelijkheden voor invoering bij de ontwikkeling van een
intersectoraal programma op 't Beeckland**

**Assessment of pre-vocational student competencies using rubrics:
possibilities for introduction while developing an
intersectoral program at 't Beeckland**

Bert Mulderij

Vorden, juni 2012



Master Leren en Innoveren

Stoas Hogeschool, Wageningen

Churchillweg 68
6706 AD Wageningen
088-0206700



Ulenhofcollege, loc. 't Beeckland

Het Hoge 41
7251 XV Vorden
0575-551155

Studentnummer: 591209001

Begeleider: L. Spit

Opdrachtgever: S. Stans

Competenties van vmbo-leerlingen beoordelen met behulp van rubrics: mogelijkheden voor invoering bij de ontwikkeling van een intersectoraal programma op 't Beeckland

Assessment of pre-vocational student competencies using rubrics: possibilities for introduction while developing an intersectoral program at 't Beeckland

Samenvatting

In dit onderzoek is een antwoord gezocht op de vraag of competentiebeoordelingen bij vmbo-leerlingen met behulp van rubrics tegelijk kunnen worden ingevoerd met de ontwikkeling van een intersectoraal programma bij Ulenhofcollege, locatie 't Beeckland. Er is geïnventariseerd wat de deelnemers aan het onderzoek al aan competentiebeoordelingen van leerlingen in hun eigen lessituatie deden voor het onderzoek van start ging. De interventie bestond uit de introductie van, en kennismaking met competentiebeoordelingen met behulp van rubrics, waarbij de leden van de experimentele groep vertrouwd zijn gemaakt met de beoordelingsmethodiek door (video-opnames van) leerlingen te beoordelen. Daarna zijn competentiebeoordelingen met deze rubrics in een eigen lessituatie uitgetest. Uit de analyse van pretest en posttest vragenlijsten blijkt dat er na drie maanden geen betrouwbaar verschil is waargenomen tussen de controlegroep en de experimentele groep. Maar het verschil in wat docenten (zowel van de beroepsgerichte als de algemeen vormende vakken) zouden willen en wat ze daadwerkelijk doen met leerlingbeoordelingen, lijkt kleiner te worden door de ervaring die ze opdoen met het werken met rubrics.

Er lijkt voldoende draagvlak te zijn om daadwerkelijk te starten met de invoering. Het is wel nodig aandacht te geven aan een aantal essentiële randvoorwaarden, zoals begeleiding, scholing en tijd.

Summary

This study answers the question whether competency assessments for pre-vocational students using rubrics can be introduced simultaneously with the development of an intersectoral program at Ulenhof College, location 't Beeckland. Before the investigation was started the participants' experience with competency assessments of students has been listed. The intervention consisted of the introduction of, and orientation to competency assessments using rubrics. The members of the experimental group were made familiar with the assessment methodology by using a video recording of the students. They tested this method of competency assessment using rubrics in their own teaching situation. Analysis of pretest and posttest questionnaires shows that after three months no reliable differences were observed between the control and experimental groups. But the difference in what teachers (both vocational and general education) would like and what they actually do with student reviews, seems to be decreasing by the experience they gain in working with rubrics.

There seems to be enough support to actually start the introduction. Attention has to be paid to some essential preconditions, such as counselling, training and time.

Inleiding

Bij de invoering van het voorbereidend middelbaar beroepsonderwijs (vmbo) per 1 augustus 1999 is het vmbo ingedeeld in sectoren met de daarbij behorende afdelingen en afdelingsprogramma's. In de laatste leerjaren van het voortgezet onderwijs moeten leerlingen in het vmbo een sector kiezen. Daarbij kan gekozen worden uit vier sectoren: Techniek, Economie, Zorg en welzijn en Landbouw. Vanaf 2007/'08 bestaan er in het vmbo naast de opleidingen in deze sectoren ook een aantal intersectorale programma's (Regeling intra- en intersectorale programma's v.m.b.o., 2008). Hierbij worden onderdelen uit twee of meer sectoren samengevoegd tot één programma. Op dit moment mogen op vmbo-scholen de volgende intersectorale programma's worden aangeboden: "Sport, dienstverlening en veiligheid", "Technologie in de Gemengde Leerweg", een ICT-programma (de "ICT-route") en een algemeen intersectoraal programma ("ISP"). Al deze programma's richten zich op jongeren die nog geen keuze voor een vmbo-sector willen maken. Leerlingen kiezen de laatste jaren in

toenemende mate voor zo'n intersectoraal programma. Die keus lijkt ten koste te gaan van de 'traditionele' sectoren (CBS, 2011). Het landbouwonderwijs, dat valt onder het ministerie EL&I, heeft tot nu toe niet deelgenomen aan de bestaande intersectorale programma's, die onder het ministerie OCW vallen. Toch zijn er inmiddels plannen om alle sectoren (dus inclusief Landbouw) in één nieuw breed intersectoraal programma onder te brengen (Van Bijsterveldt, 2011). Het eerste landelijke centrale examen van zo'n 'Intersectoraal programma nieuwe stijl' is voorzien in 2016.

Het Ulenhofcollege locatie 't Beeckland is een vmbo in Vorden met een basisberoepsgerichte leerweg (BBL), een kaderberoepsgerichte leerweg (KBL) en een gemengd/theoretische leerweg (GTL). Het is een kleinschalige school (425 leerlingen; stand 01-05-2012) met een duidelijke regiofunctie en staat goed bekend om zijn leerlingenzorg. Er wordt in de bovenbouw onderwijs gegeven in drie sectoren voor BBL en KBL: Economie (afdeling handel/verkoop), Zorg en Welzijn (afdeling Verzorging) en Landbouw (landbouwbreed; afdeling Groen); voor GTL is er een intersectoraal programma (Dienstverlening en Commercie; Codi).

Het streven van 't Beeckland is om te groeien van een 'voldoende' school naar minimaal een 'goede' of zelfs een 'excellente' school. Hierbij zal de zorg voor de leerlingen altijd voorop blijven staan. Maar de laatste jaren is de onderwijskundige ontwikkeling naar verhouding wat onderbelicht (geweest). Zo zou de transitie van basisonderwijs naar vmbo en van vmbo naar mbo sterk verbeterd kunnen worden door meer op elkaar af te stemmen. Bijvoorbeeld voor het 'plannen': bij het basisonderwijs zijn veel leerlingen gewend te werken met een (week-)taak; op het vmbo ('t Beeckland) wordt de planning voor de meeste leerlingen klassikaal (door de docent) bepaald; en op het mbo zal de leerling het weer zelf moeten kunnen (doen).

't Beeckland is in 2009 voor de GTL gestart met een intersectoraal programma. In schooljaar 2013-2014 wordt het nieuwe, brede intersectorale programma ingevoerd voor leerjaar 3 van de BBL en de KBL. Hierbij is een nadrukkelijk uitgangspunt dat de sector Landbouw bij dit programma ingepast moet worden. Vanaf 2012 zullen landelijke pilots starten voor centraal ontwikkelde examens, op weg naar wettelijke verplichting in 2014 (Onderwijsraad, 2009). Momenteel wordt op 't Beeckland sterk overwogen om aan zo'n pilot mee te doen. Met de invoering van het ISP hoopt de school op een betere concurrentiepositie in het regionale krachtenveld, en daarmee het bestaansrecht voor de komende jaren veilig te stellen.

Voor de keuze voor een intersectoraal programma zijn er drie hoofdredenen te noemen. Ten eerste is er krimp van het aantal vmbo-leerlingen te verwachten. Dit wordt deels veroorzaakt door demografische ontwikkelingen. De komende twintig jaar zullen de gemeenten in de Achterhoek te maken krijgen met teruglopende bevolkingsaantallen: de krimp zal 8 procent bedragen (Huiskamp, 2007). Maar ook zal het aantal vmbo-leerlingen ten opzichte van andere onderwijsvormen afnemen. Landelijk gezien is in de afgelopen tien jaar het aantal vmbo'ers met 8 procent afgenomen. Daar staat tegenover dat het aantal leerlingen in het praktijkonderwijs is verdubbeld en het aantal leerlingen op zowel de havo als het vwo met 29 procent is toegenomen (CBS, 2011). In 2010/'11 nam binnen het vmbo het aantal leerlingen in de beroepsgerichte leerwegen BBL en KBL opnieuw sterk af, terwijl het aantal leerlingen in de GTL vrijwel gelijk bleef. Tegelijkertijd nam het aantal leerlingen met een indicatie voor leerwegondersteunend onderwijs (lwoo) ten opzichte van 2000/'01 duidelijk toe en daalde het aantal zonder indicatie sterk (CBS, 2011).

Ten tweede betekent dit dat er een uitstel is voor de sectorkeuze van de leerling. Leerlingen die zichzelf beter kennen in relatie tot de beroepspraktijk, en die dus een beter ontwikkelde beroepsidentiteit hebben, hebben (betere) sturing over hun keuzes ten aanzien van het soort werk dat ze ambiëren. Ze weten wat ze daar in de opleidingsfeer voor moeten doen, dan wel wat ze daarvoor zelf graag willen leren (Den Boer, Mittendorff, & Sjenitzer, 2004). Daarom is het dan ook goed om leerlingen op zo vroeg mogelijke leeftijd ervaring op te laten doen met de praktijk van werk, bijvoorbeeld middels stage. Uit onderzoek bij het vmbo blijkt dat "exploratie en ervaringen" met de beroepspraktijk het beeld bepalen van de wereld, het type werk en het beroep dat vmbo-leerlingen kiezen (Den Boer et al., 2004). Voor "verwerking" van deze ervaringen blijken leerlingen meer mensen van buiten dan van binnen de school te raadplegen. Ze spreken hierover met vrienden, ouders en de stagebieder, maar relatief weinig op school, in de klas, met de docent of met medeleerlingen. Wanneer leerlingen wél met docenten over hun toekomst praten, gaat het over de vervolgopleiding die ze moeten kiezen. Identiteitsontwikkeling en bewuste keuzeprocessen zijn niet per definitie hetzelfde als een breed onderwijsaanbod. Teveel uitstel van keuze is ook niet aanbevelingswaardig. Beide veronderstellingen zijn zeer gangbaar in het vmbo. Breed aanbod heeft alleen zin als het gepaard gaat met indringende ervaringen die leerlingen zicht geven in zijn eigen mogelijkheden (Hamstra & van den Ende, 2006).

En ten derde is er een betere aansluiting en doorstroming op het mbo te verwachten. Bij leerlingen die een intersectoraal programma op het vmbo hebben gevolgd, is er minder uitval en een hoger slagingspercentage op het mbo dan bij leerlingen die een traditionele sector hebben gevolgd (Wester & van Kessel, 2012).

Naast de sectorkeus is ook een juiste determinatie van het niveau van een leerling van groot belang. Niet alleen is het voor de leerling prettig om op het juiste niveau (in de juiste leerweg) te werken, eenmaal in de bovenbouw is het vrijwel onmogelijk om qua niveau nog op te stromen. Het optimale moment van selectie is moeilijk te bepalen. Tegenover de positieve effecten van vroege selectie voor de subtop en topleerlingen staan negatieve effecten voor de vervolgloopbanen van modale leerlingen. Vroege selectie heeft een negatief effect op de vervolgloopbanen van leerlingen met een vmbo-advies (Van der Steeg, Vermeer, & Lanser, 2011).

De determinatie wordt op 't Beeckland voor een groot deel gebaseerd op basis van rapportcijfers en adviezen van docenten. Maar de mening waar een 'typische' BBL-, KBL- of GTL-leerling aan moet zou moeten voldoen blijkt vaak per docent te verschillen.

In de examenprogramma's van de beroepsgerichte vakken bij het vmbo wordt nadrukkelijk aandacht gegeven aan competenties. Van het begrip 'competentie' zijn verschillende omschrijvingen te vinden. Het is een containerbegrip dat van oorsprong meer vanuit de bedrijfskundige hoek, en minder vanuit de onderwijskundige hoek komt. In de onderwijscontext worden competenties bekeken vanuit hun ontwikkelbaarheid en hun beïnvloedbaarheid (Valcke, 2010). Sinke (2006) gaf (op basis van een definitie van Parry uit 1996) de volgende beschrijving van een competentie:

[Een competentie is] een cluster van verwante kennis, vaardigheden en houdingen die van invloed is op een belangrijk deel van iemands taak (een rol of verantwoordelijkheid), die samengaat met de prestatie op de taak, die kan worden gemeten en getoetst aan aanvaarde normen, en die kan worden verbeterd door middel van training en ontwikkeling.

(Sinke, 2006, p.42)

Valcke (2010) voegt daar nog aan toe dat dat plaatsvindt in concrete, complexe en authentieke situaties.

In het vmbo gaat het bij competenties om situaties die samenhangen met een bepaald beroep en/of functie. De leerling voert handelingen uit die in de gegeven situatie het meest passend - adequaat - zijn. Een competentie in het vmbo is derhalve een combinatie van kennis, vaardigheden, houdingen en persoonlijke ontwikkeling (De Jong, Hoekstra, van der Wel, & Wensink, 2005).

Kerncompetenties zijn competenties die, bij wijze van spreken, iedereen overal nodig heeft. Niet alleen op school, maar ook op de werkplek, in de samenleving en in de privésfeer. De toepassing van deze kerncompetenties is niet beperkt tot een bepaalde vakdiscipline, beroepsgroep of sector. Kerncompetenties overstijgen het niveau van individuele vakken of leertaken. Ze zouden als een rode draad door het curriculum van elke opleiding moeten lopen (De Boer & Kuhlemeier, 2005).

Iedere schakel in de onderwijskolom vmbo-mbo-hbo-wo definieert de groei en ontwikkeling van zijn student volgens een eigen methodiek in een eigen taal. Leerlingen en studenten kunnen hun voorkennis verderop in de kolom moeilijk verzilveren. Deze "onvergelijkbaarheid" bemoeilijkt een soepele doorstroom van studenten in de onderwijskolom. Daarom is er door docenten van vmbo-, mbo-, hbo-scholen en Wageningen UR, onder leiding van de projectgroep Ruggengraat van de Groene Kennis Coöperatie, één taal gedefinieerd waarin de groei en ontwikkeling van studenten in de hele kolom helder en eenduidig beschreven kan worden. Hoewel het een samenwerkingsproduct betreft van het groene onderwijs is 4Cyourway in de volle breedte van het onderwijs inzetbaar (<http://www.4cyourway.nl/>). Voor alle ingrediënten van opleidingen, zoals algemene competenties, vakinhoudelijke onderdelen en andere uitstroomseisen (bijvoorbeeld taal en exact), is een gemeenschappelijke maatstaf geformuleerd. Op de schaal van deze eigen maatstaf van de competentie of de vakinhoud is vervolgens per schakel van de kolom de hoogste uitstroomseis gedefinieerd. Daarvan zijn in een later stadium de onderliggende groeistappen afgeleid (rubrics), die zoveel mogelijk in termen van waarneembaar studentengedrag zijn beschreven (De Jong, 2008).

Voor het beoordelen van competenties zal er op één of andere manier 'getoetst' moeten worden. In de literatuur over competentiegerichtoetsen wordt vaak het begrip *toets* tegenover het begrip *assessment* geplaatst om een verschil aan te duiden tussen klassieke (meerkeuze)toetsen en nieuwe vormen van beoordelen. Met *assessment* worden dan nieuwe methoden bedoeld. De laatste jaren zijn er veel nieuwe methoden ontwikkeld en in gebruik genomen, zoals portfolio's en proeven van bekwaamheid. Deze nieuwe methoden zijn echter geen vervanging voor bestaande methoden. Competenties zijn complexe gehelen van kennis, vaardigheden en attitudes, en dit maakt het lastig om vast te stellen in hoeverre een leerling een competentie heeft verworven. Het is tevens praktisch onmogelijk om met één enkele methode een competentie te beoordelen (Baartman, van der Sanden, Seezink, & van der Brandt, 2006). Een mogelijke oplossing is het gebruik van een Competentie Assessment Programma (CAP). Hierbij wordt een combinatie van verschillende methoden

gebruikt, bestaande uit zowel klassieke toetsen als nieuwe assessmentmethoden (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2006). Het bijbehorende kwaliteitskader omvat dan ook zowel traditionele kwaliteitscriteria als criteria die beter aansluiten bij de beoordelingsmethode. Voorbeelden van kwaliteitscriteria zijn: Authenticiteit (overeenkomst met de beroepspraktijk), Cognitieve complexiteit (beoordelen van cognitieve vaardigheden nodig voor beroep), Eerlijkheid (geen bevoor- of benadeling), Transparantie (duidelijke en begrijpelijke criteria en procedures), Directheid (geen vertaling van gedrag in “CAP-situatie” naar praktijk nodig), Herhaalbaarheid (verschillende situaties of beoordelaars), Vergelijkbaarheid (voor alle deelnemers gelijk, consistent) en Tijd en kosten (haalbaarheid uitvoering) (Dijkstra & Baartman, 2011).

Het beoordelen van de leerlingen van het Ulenhofcollege locatie 't Beeckland is tot op heden voornamelijk gebaseerd op cognitieve vaardigheden. Ook in de missie van de school wordt voornamelijk over leren gesproken in de vorm van kennis opdoen (Ulenhofcollege, 2011). Inmiddels is er in de concepten voor een nieuw schoolbeleidsplan wel meer oog voor de ontwikkeling en de omgeving van de leerling. Ook docenten van 't Beeckland, vooral van de beroepsgerichte vakken, gaan uit van een andere, bredere kijk op het leren en willen zich van daar uit graag meer met de ontwikkeling van competenties van leerlingen bezighouden. Leren is meer dan alleen kennis tot je nemen. Leren doet iedereen zelf, maar zoals Vygotski aangaf, is de sociale context daarbij ook van groot belang (Valcke, 2010). Hieruit voortvloeiend zouden er twee benaderingen tegenover elkaar gezet kunnen worden: ‘independent learning’ en ‘dependent learning’ (Kirschner (2006). ‘Independent learning’ is gebaseerd op een competitieve en individualistische basis. Leren, vaak gezien als het opstapelen van kennis en vaardigheden, wordt beschouwd als een individuele verwezenlijking van iedere leerling zelf (en wordt als zodanig beoordeeld). ‘Dependent learning’ is leren waarbij de leerdoelen, de leerprocessen worden gebruikt. Uiteindelijk zal de beoordeling, het assessment geheel afhangen van anderen. Leren is hier meer dan het opdoen van kennis en vaardigheden. Het is meer het opdoen van competenties, die niet alleen individuele kennis en cognitieve vaardigheden omvat, maar ook houding en interpersoonlijke vaardigheden (bijvoorbeeld sociale en emotionele). Het leerproces is dan veel meer een groepsproces dat leidt tot een groepsresultaat.

Kirschner (2006) geeft aan dat we geneigd zijn steeds keuzes te maken (‘dichotomie’), dus kiezen we óf voor independent learning, óf voor dependent learning. Hij pleit er voor deze twee benaderingen te integreren en een tussenvorm te ontwikkelen: ‘interdependent learning’.

Illeris (2002) heeft juist niet gekozen om deze twee benaderingen te integreren. Hij geeft aan dat leren het geheel is van processen die leiden naar permanente verandering van capaciteit (of dat nou fysiek, cognitief, emotioneel of sociaal van aard is), en dat dit niet exclusief afhankelijk is van biologische ontwikkeling of veroudering zoals Piaget die beschrijft (Valcke, 2010). Dit houdt volgens Illeris (2002) in dat leren bepaald wordt door structuur met twee verschillende typen processen en drie dimensies. Aan de ene kant zijn er interactieprocessen tussen de leerling en de omgeving, aan de andere kant zijn er de innerlijke mentale processen. De drie dimensies zijn de cognitieve, de emotionele en de sociale dimensies.

Leren is dus niet alleen het verwerken van theoretische informatie. Het is ook leren door ervaring, door sociale interactie en door kritische zelfreflectie. Daarom zullen deze aspecten in goed (competentiegericht) onderwijs structureel aandacht moeten krijgen.

Bij gedifferentieerd leren zullen leerlingen niet meer op hetzelfde moment met dezelfde beoordelingscriteria en op hetzelfde niveau aan de slag gaan (Sluijsmans, 2008). Ze volgen een individueel leerpad met een daarop aansluitende beoordeling. Sluijsmans (2008) geeft aan dat die beoordeling met een ipsatief beoordelingsmodel gedaan zou kunnen worden. Beoordelingen van leerlingen worden dan niet vergeleken met het gemiddelde van hun klas of met vooraf vastgestelde beoordelingscriteria en standaarden, maar worden individuele beoordelingen die zijn aangepast naar het niveau van de leerling vergeleken in de tijd. Dat kan door aan (clusters van) beoordelingscriteria standaarden te koppelen (scoringsrubrieken) die worden vastgesteld door de moeilijkheid van de taak. Elke leerling bouwt een individueel portfolio op dat inzichtelijk maakt op welke beoordelingscriteria vooruitgang wordt geboekt. Zo’n ipsatieve beoordeling is gericht op toekomstig leren en wordt gestuurd door de vraag hoe de leerling verder kan groeien (Ward, Gruppen, & Regehr, 2002).

Onder een rubric (scoringsrubriek) wordt een descriptieve beoordelingsschaal verstaan, waarin de schaalpunten voorzien zijn van omschrijvingen (indicatoren; voorbeelden) van de handelingen, werkwijzen, producten en/of houdingen die indicatief zijn voor het desbetreffende ‘niveau’. De beoordelaar beoordeelt de prestatie van de kandidaat door deze te vergelijken met de beschrijvingen van de beoordelingsschaal (De Boer & Kuhlemeier, 2005). Gulikers (2008) geeft de volgende definitie van een rubric: “een beschrijving van een aantal concreet te observeren gedragsniveaus (van eenvoudiger naar complexer) op een beoordelingscriterium”

Jonsson en Svingby (2007) geven aan dat door het gebruik van rubrics een betrouwbare score (van prestaties, competenties) verhoogd kan worden, maar ook dat rubrics niet persé tot een valide beoordeling hoeven te leiden. Verder geven ze aan dat rubrics de potentie hebben om het leren te bevorderen en/of de

instructie te verbeteren. Een bezwaar van het gebruik van rubrics zou kunnen zijn dat ze op leerlingen verlammend zouden kunnen werken en de creativiteit wegnemen (Cooper & Gargan, 2009). Rubrics kunnen te sturend werken, waardoor iedereen wordt aangezet om op dezelfde manier naar problemen en oplossingen te kijken. 'Outside the box'-denken wordt daardoor ontmoedigd.

Competentiegericht onderwijs betekent niet alleen dat men het onderwijs in de klas gaat vernieuwen, maar dat de hele organisatie moet worden aangepast (Dochy & Nickmans, 2005). Zo zijn de toetsen in competentiegericht onderwijs niet meer leerwegafhankelijk en leerstofgericht, zoals in het traditionele onderwijs, maar is de assessment leerwegaafhankelijk en context- en competentiegericht. In plaats van vaste toetsmomenten met toetsen die door de docent worden beoordeeld komen er flexibele assessmentmomenten, waar door onafhankelijke assessoren wordt beoordeeld. En in plaats van examens die als sluitstuk fungeren, is assessment ook een startpunt voor het ontwerp van flexibele leerroutes. Kortom: het leerwegafhankelijke, aanbodgerichte curriculum uit het traditionele onderwijs verandert in een vraaggestuurd curriculum met een stijgende zelfsturing in het competentiegerichte onderwijs (Dochy & Nickmans, 2005)

Voor een succesvolle innovatiestrategie is het van belang rekening te houden met de stadia van betrokkenheid. Betekenisvolle contexten, waarbij een link wordt gelegd tussen vernieuwingen en de 'eigenheid' van de leraar, bevorderen dat leraren weerstanden overwinnen en het praktische nut van een vernieuwing ervaren. Methoden om dit te bevorderen zijn reflectie en werkplekleren, waarbij de eigen beroepspraktijk van de leraar centraal staat (Van der Bolt, Studulski, van der Vegt, & Bontje, 2005). Om eigenaarschap bij docenten te creëren kan met een kleine groep docenten gestart worden, om daarna het aantal daarna uit te breiden. Een andere benadering zou kunnen zijn om alle docenten uit de organisatie vanaf de start er bij betrekken. Onderzoeksresultaten geven aan dat breed beginnen tot een trager veranderingsproces leidt, maar dat de kans groter is dat alle docenten zich op termijn bij het proces betrokken gaan voelen (Waslander, 2007). Klein beginnen lijkt een snellere start op te leveren, maar heeft ook het risico dat andere docenten zich op langere termijn blijven verzetten.

"Het paradoxale is dat de strategie om klein te beginnen meer kans van slagen heeft in scholen met een sterk sociaal weefsel, terwijl de noodzaak voor een dergelijke strategie uitgerekend in die scholen minder groot is. Scholen met een zwak sociaal weefsel zien zich vaak genooddaakt klein te beginnen, terwijl dat juist daar met meer risico's gepaard gaat."
(Waslander, 2007, p.53)

Vernieuwing komt er op neer dat docenten iets anders gaan doen dan wat ze daarvoor deden. Van een aanpak die niet leidt tot ander gedrag van docenten, valt geen enkel effect te verwachten (Wikeley, Stoll, Murillo, & de Jong, 2005). Door bij veranderingsprocessen 'positive psychology' te gebruiken, wordt niet uitgegaan van tekortkomingen bij mensen, maar juist van hun kracht (Seligman & Csikszentmihalyi, 2000) en zal de kans op gedragsverandering vergroot worden. 'Bottom-up' vernieuwen, dus werken vanuit de wensen en behoeften van de docent, werkt beter dan 'top-down', waarbij gewerkt wordt vanuit het zogenoemde 'expertmodel' (Bergen & van Veen, 2004). Daarbij ontwikkelen experts en wetenschappers een innovatie en bedenken de implementatie, de inhoud en de scholing. Er wordt zo weinig rekening gehouden met de beleving van de leerkracht en psychologische factoren. Veel van deze innovaties zijn dan ook weinig succesvol gebleken (Van der Bolt, Studulski, van der Vegt, & Bontje, 2005). Bij 'bottom-up' wordt gewerkt vanuit het 'craft-model'. Hierin staat de praktijkkennis van de docent centraal. De praktische aspecten van het lesgeven dienen als basis voor professionele ontwikkeling en vernieuwing (Bergen & van Veen, 2004).

Veranderen en innovatie is niet alleen realiseren van grootse plannen, maar verloopt stapsgewijs. Van den Berg en Vernooy (2000) geven aan dat vernieuwen "in haar beste vorm een evolutionair proces [is] dat van binnenuit, stap voor stap wordt ontwikkeld en langzaam maar zeker geïntegreerd raakt in het functioneren van de school" (p. 38) .

De invoering van het nieuwe intersectorale programma op 't Beeckland is in schooljaar 2011/12 gestart door aan een onderwijsontwikkelgroep de opdracht te geven dit voor te bereiden en uit te voeren. De veranderingen in de bovenbouw zullen enorm zijn: behalve vakinhoudelijk zal ook het totale vakkenpakket veranderen en zullen organisatorische aanpassingen nodig zijn. Het ligt voor de hand om ook te bekijken of er nog andere aspecten zijn die tegelijk kunnen worden aangepakt, zodat er maar één keer een grote verandering in de organisatie hoeft plaats te vinden. Het inbedden in het programma van structurele aandacht voor het beoordelen van competenties van leerlingen is zo'n aspect. Daarom is dit onderzoek uitgevoerd. De uitvoering heeft plaats gevonden met een deel van het personeel: de leden van het team Basisberoepsgerichte leerweg Bovenbouw, één van de vier teams die 't Beeckland kent, aangevuld met de leden van de onderwijsontwikkelgroep die zich met de invoering van het intersectorale programma bezighouden.

De centrale vraag voor dit onderzoek is: Zijn competentiebeoordelingen bij leerlingen met behulp van rubrics structureel te gebruiken op 't Beeckland vanaf het moment dat het intersectorale programma ingevoerd wordt? Er is geïnventariseerd wat de deelnemers aan het onderzoek al aan competentiebeoordelingen van leerlingen doen in hun eigen lessituatie. De deelnemers zijn vertrouwd gemaakt met een methodiek door (video-opnames) van leerlingen te beoordelen aan de hand van rubrics. Hierbij is er gekozen om te werken met rubrics die onder begeleiding van het CPS zijn ontwikkeld in het project De Groene Werkplek, dat uitgevoerd is op zes vmbo-scholen (Nijman, 2008). Het examenprogramma vmbo-groen hanteert - net als het mbo - de zogenaamde SHL-competenties. Er zijn 25 SHL-competenties, waarvan het vmbo er 18 gebruikt. De meeste van deze SHL-competenties zijn ook in de andere sectoren goed bruikbaar, zoals bijvoorbeeld 'beslissen/activiteiten initiëren', 'samenwerken en overleggen', 'ethisch en integer handelen', 'relaties opbouwen' en 'presenteren'. In een leerlingentool zijn alle niveaus (BBL, KBL en GTL) heel begrijpelijk en herkenbaar beschreven met korte gedragsomschrijvingen, waarmee de leerling kan bepalen in welke mate de afzonderlijke competenties zijn ontwikkeld (Van 't Wout & de Jong, 2010). Er zijn vier rubrieën: 'starter', 'geoefend', 'gevorderd' en 'op niveau'. Bij de competentie 'samenwerken en overleggen' kan bijvoorbeeld worden aangegeven "ik voer mijn taken op eigen initiatief uit zonder overleg met anderen" (niveau 'starter') of "ik waardeer de bijdragen van anderen en lever een bijdrage aan de goede sfeer in de groep" ('op niveau').

Aan alle leden van de experimentele groep is gevraagd om een competentiebeoordeling met deze rubrics in een lessituatie uit te proberen.

Er is onderzocht of het verschil in wat docenten willen en wat ze daadwerkelijk doen omtrent leerlingbeoordelingen verandert wanneer ze ervaring opdoen met het werken met rubrics. Hierbij zijn met name de kwaliteitscriteria 'eerlijkheid', 'transparantie', 'directheid', 'herhaalbaarheid' en 'vergelijkbaarheid' bestudeerd. Deze kwaliteitscriteria zijn afkomstig uit het kwaliteitskader van Baartman, Bastiaens, Kirschner, en Geurts (2005). Ook is gekeken of het daarbij uitmaakt of de docent een beroepsgericht vak geeft of een algemeen vormend vak.

Tot slot is gevraagd naar aanbevelingen voor het vervolg: hoe kunnen we als team en/of organisatie verder met de opgedane ervaringen? In hoeverre kan de – middels dit onderzoek - ontwikkelde kennis en opgedane ervaring een basis vormen voor de werkwijze bij de in te voeren intersectorale programma's?

Methoden van onderzoek

Dit onderzoek betreft een quasi experimenteel pre-test post-test design. Aangezien de experimentele groep niet random was samengesteld, is er sprake van een Non-Equivalent Groups Design (Trochim, 2006).

De pretest was een schriftelijke enquête met open en gesloten vragen. De interventie met de experimentele groep bestond uit de introductie van, en kennismaking met competentiebeoordelingen met behulp van rubrics, gevolgd door het in een eigen lessituatie toepassen van een competentiebeoordeling bij leerlingen met rubrics. De posttest was weer een schriftelijke enquête met open en gesloten vragen.

Onderzoekseenheden

Het onderzoek is uitgevoerd met 17 onderwijsgeevenden die werkzaam zijn op het Ulenhofcollege, locatie 't Beeckland te Vorden. Daar werkten op 01-11-2011 in totaal 47 onderwijsgeevenden bij ofwel een algemeen vormend vak (avo, 35 docenten), ofwel een beroepsgerichte vak (bgv, 12 docenten).

De experimentele groep bestond uit 9 personen: 5 avo-docenten (gemiddelde leeftijd 52.8 jaar, $SD = 4.9$; 3 ♀ en 2 ♂) en 4 bgv-docenten (45.5 jaar, $SD = 16.4$; 3 ♀ en 1 ♂). Zij zijn lid van het team 'Basisberoepsgerichte Leerweg - Bovenbouw' en van de onderwijsontwikkelgroep 'invoering intersectoraal programma'.

De controlegroep bestond uit 8 personen: 5 avo-docenten (46.4 jaar, $SD = 16.5$; 3 ♀ en 2 ♂) en 3 bgv-docenten (54.0 jaar, $SD = 4.6$; 2 ♀ en 1 ♂). Dit zijn personen die gereageerd hebben op de eerste enquête.

Meetinstrumenten

In de open vragen van de pretest is gevraagd of de respondent in zijn/haar lessen competenties van leerlingen heeft beoordeeld. Bij het antwoord 'ja' werd verzocht om te beschrijven welke competenties werden beoordeeld, hoe dat plaats vond en welk materiaal daar eventueel bij gebruikt is (bijvoorbeeld scorelijsten).

De gesloten vragen van pretest en posttest zijn gebaseerd op de vragenlijst uit het onderzoek naar kwaliteitscriteria voor CAP's (in het mbo) van Baartman, Bastiaens, Kirschner, en Geurts (2005). Uit deze vragenlijst zijn een aantal stellingen genomen die (ook) van toepassing zijn op het vmbo. De gebruikte stellingen gaan grotendeels over beoordelen in het algemeen, maar door deze vragen te plaatsen in de context van competentiebeoordeling werden de respondenten gedwongen hier bewust over na te denken.

Competentiebeoordelingen met rubrics op 't Beeckland

De gebruikte kwaliteitscriteria en de bijbehorende stellingen staan in tabel 1. Bij elke stelling zijn steeds twee vragen gesteld: “Hoe belangrijk vind je dit bij een competentiebeoordeling?” (= A-vraag) en “In hoeverre vind je dat je dit hebt gerealiseerd?” (= B-vraag). De respondenten gaven dit aan op een 5 punts-Likertschaal. Deze schaal varieerde van ‘1 = niet’, via ‘2 = nauwelijks’, ‘3 = neutraal’, en ‘4 = enigszins’, naar ‘5 = wel’. Bij de pretest en de posttest zijn dezelfde 13 stellingen en bijbehorende vragen gebruikt.

De deelnemers uit de experimentele groep hebben bij de posttest behalve de gesloten vragen ook open vragen beantwoord. Daarin is gevraagd naar de ervaringen met competentiebeoordelingen van leerlingen in de eigen lessituatie, naar de mogelijkheden en/of beperkingen die daarbij werden gezien, naar de eventuele toegevoegde waarde van de tijdens het experiment gebruikte rubrics, en naar waar vooral op gelet moet worden als er standaard meer aandacht gegeven gaat worden aan competentiebeoordelingen van leerlingen. Tot slot is de vraag gesteld of competentiebeoordelingen van leerlingen met behulp van rubrics op 't Beeckland structureel toegepast zouden kunnen worden.

Tabel 1. In pretest en posttest gebruikte kwaliteitscriteria en bijbehorende stellingen. Bij elke stelling zijn steeds twee vragen gesteld: A - “Hoe belangrijk vind je dit bij een competentiebeoordeling?” en B - “In hoeverre vind je dat je dit hebt gerealiseerd?” (Bron: Baartman, Bastiaens, Kirschner, & Geurts, 2005)

Kwaliteitscriterium	Stelling
Eerlijkheid	1. De beoordelingsopdracht sluit aan bij het niveau van de individuele leerling.
	2. Er is vooraf bepaald welke onderdelen belangrijk zijn en zwaarder meetellen in de eindbeoordeling.
	3. Subjectieve of niet in de criteria opgenomen zaken tellen niet mee voor de beoordeling.
Transparantie	4. Leerlingen weten vooraf of de beoordeling meetelt voor het rapport/schoolexamen.
	5. Voor leerlingen is vooraf duidelijk wat de beoordelingscriteria zijn.
	6. Voor leerlingen is het duidelijk wat de beoordelingsprocedure is.
Directheid	7. De beoordelingscriteria zijn gericht op direct waarneembare zaken.
	8. Door de manier van beoordelen blijkt direct hoe een leerling zal functioneren in een beroep.
Herhaalbaarheid	9. Als twee personen de opdracht beoordelen (tegelijkertijd en/of in verschillende klassen), wordt gecontroleerd of hun beoordelingen overeenkomen.
	10. Als de beoordeling herhaald wordt, wordt gecontroleerd of dit dezelfde uitkomst oplevert.
Vergelijkbaarheid	11. De beoordelingsopdrachten zijn voor alle leerlingen gelijk.
	12. De beoordelingscriteria zijn voor alle leerlingen gelijk.
	13. Alle leerlingen worden in een vergelijkbare situatie beoordeeld.

Interventie

In januari 2012 is aan de leden van de experimentele groep in een workshop van 100 min. uitgelegd wat rubrics zijn en hoe je ze zou kunnen gebruiken, om vervolgens te oefenen met de Groene Werkplek/CPS-rubrics aan de hand van video-opnames van leerlingen. Deze video-opnames zijn in december 2011 gemaakt van leerlingen (leerjaar 4; BBL en KBL; afdeling Groen) die vertelden (presenteerden) wat zij in de voorafgaande weken in de praktijklessen hebben gedaan.

Aan het begin van de workshop is de opname van leerling 1 zonder inhoudelijke inleiding getoond met alleen de vraag: hoe beoordeel je deze leerling? Na het delen van de opmerkingen is leerling 2 getoond en besproken. Daarna zijn de rubrics van De Groene Werkplek/CPS geïntroduceerd en toegelicht. Met behulp van de rubrics behorende bij de competentie ‘presenteren’ (afbeelding 1) zijn vervolgens de leerlingen 3 en 4 bekeken en besproken.

De deelnemers aan de workshops hebben de opdracht meegekregen om in februari en maart in hun eigen lessituatie een keer een beoordeling te doen met behulp van de Groene Werkplek/CPS-rubrics voor de competenties ‘presenteren’, ‘samenwerken en overleggen’ of ‘formuleren en rapporteren’. Gedurende deze maanden is de uitvoering van de opdracht geagendeerd en besproken in zowel de teamvergaderingen als de bijeenkomsten van de onderwijsontwikkelgroep. Daar werden de ervaringen uitgewisseld, vragen gesteld/beantwoord en de resultaten besproken.

Competentie	PRESENTEREN © CPS			
leerweg	Starter	Geoefend	Gevorderd	Streefniveau
BBL	Ik vertel, meestal goed verstaanbaar, een aantal onderwerpen.	Ik vertel een aantal onderwerpen in een vaste volgorde. Ik spreek duidelijk en kijk de luisteraars aan.	Ik vertel een duidelijk en goed verstaanbaar verhaal. Ik kijk de luisteraars aan en gebruik soms materialen om de presentatie aantrekkelijk te maken.	ik houd een duidelijk en aantrekkelijk verhaal. Ik maak contact met luisteraars om dit te controleren.
KBL	Ik vertel, meestal goed verstaanbaar, een aantal onderwerpen.	Ik vertel een aantal onderwerpen in een vaste volgorde. Ik spreek duidelijk en kijk de luisteraars aan en gebruik materialen om de presentatie aantrekkelijk te maken.	Ik vertel een duidelijk opgebouwd, aantrekkelijk en goed verstaanbaar verhaal. Ik controleer dit bij de luisteraars en vul mijn verhaal zo nodig aan.	<i>Ik vertel een duidelijk opgebouwd, aantrekkelijk en goed verstaanbaar verhaal. Ik controleer dit bij de luisteraars, vul zo nodig aan en beantwoord eenvoudige vragen. Ik geef aan waarin ik nog kan verbeteren.</i>
GTL	Ik vertel een aantal onderwerpen in een vaste volgorde. Ik spreek duidelijk en kijk de luisteraars aan en gebruik materialen om de presentatie aantrekkelijk te maken.	Ik vertel een duidelijk opgebouwd, aantrekkelijk en goed verstaanbaar verhaal. Ik controleer dit bij de luisteraars en vul mijn verhaal zo nodig aan.	Ik vertel een duidelijk opgebouwd, aantrekkelijk en goed verstaanbaar verhaal. Ik controleer dit bij de luisteraars, vul zo nodig aan en beantwoord eenvoudige vragen. Ik geef aan waarin ik nog kan verbeteren.	Ik vertel een duidelijk opgebouwd en goed verstaanbaar verhaal. Ik controleer dit bij de luisteraars en kan argumenten en antwoorden geven. Ik presenteer in verschillende situaties en geef aan waarin mijn team en ook ik zelf kan verbeteren

Abbeelding 1. Een voorbeeld van rubrics voor vmbo: de competentie 'presenteren'
(Bron: Project De Groene Werkplek / CPS).

Procedure van data verzamelen

De pretest-enquête (bijlage 1), met zowel open als gesloten vragen, is uitgezet op 03-11-2011 bij alle onderwijsgevende personeelsleden van 't Beeckland (N=46). Van het team 'basisberoepsgerichte leerweg bovenbouw' hebben 6 (van de 11) en van de onderwijsontwikkelgroep 6 (van de 7) personen de enquête ingevuld. Drie van hen zitten zowel in het team 'basisberoepsgerichte leerweg bovenbouw' als in de onderwijsontwikkelgroep, waardoor de grootte van de experimentele groep 9 personen bedraagt. De controlegroep is samengesteld uit onderwijsgevend personeel dat deze eerste enquête heeft ingevuld (8 uit 31).

Op 05-04-2012 is de posttest-vragenlijst (bijlage 2) uitgereikt aan alle 17 deelnemers aan de pretest.

Beide enquêtes zijn op naam uitgezet. Nadat de gegevens van de pretest en de posttest per deelnemer zijn samengevoegd, zijn de gegevens geanonimiseerd.

Data-analyse

De antwoorden op de open vragen uit de pretest en de posttest zijn daar waar mogelijk open gecodeerd (Boeije, 2005). Sommige respondenten hebben echter soms zulke korte antwoorden gegeven dat het antwoord al vrijwel geheel uit codes bestond. Vervolgens zijn de antwoorden per vraag geclusterd en is het resultaat in één of meer zinnen samengevat.

Om de interne consistentie van de gebruikte schalen te bepalen is een betrouwbaarheidsanalyse met behulp van Cronbach's α verricht. Hiervoor zijn de vragen per kwaliteitscriterium geclusterd. Omdat het hier om gepaarde waarnemingen gaat (repeated measures design), is de Cronbach's α zowel voor de pretest als voor de posttest berekend. Field (2009) geeft aan dat de minimale waarde van Cronbach's $\alpha > .7$ moet zijn voor een voldoende interne consistentie. Voor complexe begrippen geldt een minimale vereiste Cronbach's $\alpha > .60$ (Baarda, de Goede, & van Dijkum, 2007). Verder is de minimale waarde sterk afhankelijk van het aantal items en waarnemingen dat in de schaal is opgenomen: hoe meer items en/of waarnemingen, hoe groter de waarde van Cronbach's α wordt (Wells & Wollack, 2003). In dit onderzoek is zowel het aantal items als het aantal waarnemingen klein. Daarom is Cronbach's $\alpha > .50$ als minimale waarde voor de interne consistentie van de vragenlijst aangehouden.

Vervolgens is een tweevoudige variantieanalyse voor gepaarde waarnemingen uitgevoerd (Factorial Repeated-Measures ANOVA; Field, 2009), met als variabelen: Groep (controlegroep en experimentele groep) en Vak (bgv en avo). De analyses zijn uitgevoerd op a) de scores van alle vragen binnen de kwaliteitscriteria, b) op alleen de A-vragen, c) op alleen de B-vragen en d) op het verschil tussen de A-vraag en de B-vraag.

Resultaten

In de pretest-enquête zijn een aantal open vragen gesteld over competentiebeoordelingen. Op de vraag 'Beoordeel je in jouw lessen competenties van leerlingen' beantwoordden 13 van de 17 docenten dit met 'ja' (76.4%; 7 avo- en 6 bgv-docenten). Veruit de meest genoemde competenties zijn 'samenwerken' (10 keer) en 'presenteren' (7 keer). Andere genoemde competenties zijn 'toepassen van kennis' (4 keer: alleen door avo-docenten genoemd), 'systematisch werken' (2 keer), 'onderzoeken' (2 keer), 'instructies opvolgen' en het 'verwoorden van een mening' (beide 1 keer).

De beoordeling van deze genoemde competenties wordt door de docenten op verschillende manieren uitgevoerd. Twee docenten schreven dat zij van tevoren aangegeven aan welke eisen de leerlingen moeten voldoen. Vijf docenten gaven aan dat de klas of de leerling bij de beoordeling betrokken wordt. Ook is genoemd dat de competentiebeoordeling plaatsvindt door een schriftelijke toets en één docent gaf aan dat hij aan het einde van een lesperiode de leerling een globale beoordeling geeft in de vorm van een cijfer. Er wordt door de docenten gebruik gemaakt van scorelijsten die zelf ontwikkeld zijn of uit een lesmethode komen. Voor de normering werd door één docent genoemd dat er vooral naar exameneisen wordt gekeken. Een ander geeft aan dat er geen scorelijst is, maar dat er een punt aftrek volgt als er door de leerling niet aan de systematiek (lees 'gedrag') gehouden wordt.

Van de vier docenten die nu niet aan competentiebeoordelingen van leerlingen doen gaven er drie aan dat eventueel wel te willen, bijvoorbeeld voor de competentie 'samenwerken', maar dat dat niet ten koste mag gaan van de kennis(overdracht).

De resultaten van de analyse van de interne consistentie van de gebruikte vragenlijsten staan weergegeven in tabel 2. In de pretest waren de kwaliteitscriteria Eerlijkheid en Directheid onvoldoende consistent (Cronbach's $\alpha < .50$). De Cronbach's α van Directheid (= .19) lag erg ver onder de betrouwbaarheidsgrens. Daarom is met name dit kwaliteitscriterium niet meer in de verdere analyses betrokken.

Bij de posttest was de interne consistentie voor alle kwaliteitscriteria voldoende (Cronbach's $\alpha > .66$).

Tabel 2. De interne consistentie van de gesloten vragen uit de pretest en posttest-vragenlijst (N=17).

Kwaliteitscriterium	Pretest			Posttest		
	Cronbach's α	M	SD	Cronbach's α	M	SD
Eerlijkheid	.49	3.78	1.08	.67	3.90	0.92
Transparantie	.53	3.89	1.75	.73	4.20	1.13
Directheid	.19	3.33	1.39	.66	3.49	0.86
Herhaalbaarheid	.69	3.02	1.68	.66	3.41	1.60
Vergelijkbaarheid	.74	4.06	1.05	.74	4.28	0.95

Tabel 3. Scores op de vragenlijst voor controlegroep (N = 8) en experimentele groep (N = 9)

Kwaliteitscriterium	Vraag	Pretest				Posttest			
		Controlegroep		Exp. groep		Controlegroep		Exp. groep	
		M	SD	M	SD	M	SD	M	SD
Eerlijkheid	A	4.21	0.43	4.00	0.72	4.25	0.68	4.04	0.77
	B	3.58	0.43	3.37	1.09	3.38	0.50	3.51	0.67
Transparantie	A	4.13	0.59	4.30	0.70	4.46	0.35	4.33	0.99
	B	3.63	0.84	3.51	1.31	4.17	0.50	3.85	0.94
Herhaalbaarheid	A	3.50	0.76	3.50	1.25	4.19	0.80	3.83	1.17
	B	2.56	0.98	2.50	1.25	3.31	1.00	2.39	0.86
Vergelijkbaarheid	A	3.63	0.81	4.48	0.63	4.21	0.97	4.52	0.53
	B	3.67	0.71	4.37	0.90	4.08	0.68	4.30	0.70

Tabel 4. Scores op de vragenlijst voor bgv-docenten (N = 7) en avo-docenten (N = 10)

Kwaliteitscriterium	Vraag	Pretest				Posttest			
		Bgv		Avo		Bgv		Avo	
		M	SD	M	SD	M	SD	M	SD
Eerlijkheid	A	4.24	0.63	4.00	0.59	4.57	0.57	3.83	0.67
	B	3.48	0.77	3.47	0.91	3.90	0.46	3.50	0.65
Transparantie	A	3.95	0.65	4.40	0.58	4.71	0.41	4.17	0.85
	B	3.67	1.05	3.50	1.16	4.33	0.58	3.77	0.82
Herhaalbaarheid	A	3.64	0.90	3.40	1.13	4.43	0.53	3.70	1.16
	B	2.42	1.33	2.60	0.97	3.14	0.75	2.60	1.15
Vergelijkbaarheid	A	4.14	0.67	4.03	0.95	4.62	0.52	4.20	0.88
	B	4.67	0.58	3.60	0.78	4.43	0.46	4.03	0.78

In tabel 3 en 4 staan de resultaten van de gesloten vragen uit de enquêtes uit de pretest en de posttest weergegeven.

Bij bijna alle A-vragen (“Hoe belangrijk vind je dit?”) was de absolute score hoger dan de absolute score bij de B-vragen (“In hoeverre heb je dit gerealiseerd?”). Alleen bij het kwaliteitscriterium Vergelijkbaarheid was dit in de pretest bij avo-docenten omgekeerd. Deze verschillen zijn echter nergens betrouwbaar ($p < .10$).

Ten opzichte van de pretest zijn de scores op alle A-vragen in de posttest toegenomen bij de kwaliteitscriteria Eerlijkheid, Herhaalbaarheid en Vergelijkbaarheid, maar alleen bij Vergelijkbaarheid is deze trend in de analyse zichtbaar geworden ($p=.067$, $F(1,13)=4.01$, $\eta^2=.236$). Ook bij de B-vragen nam de score overall toe, maar alleen bij het kwaliteitscriterium Transparantie was de toename betrouwbaar ($p=.018$, $F(1,13)=7.29$, $\eta^2=.359$) en bij Eerlijkheid werd de trend in de analyse zichtbaar ($p=.087$, $F(1,13)=3.43$, $\eta^2=.209$).

Uit de analyses blijkt dat er bij de variabele Groep (controlegroep en experimentele groep) geen betrouwbare verschillen ($p < .05$) of zelfs aanwijzingen voor verschillen ($p < .10$) tussen de twee groepen zijn.

Bij de variabele Vak (bgv en avo) zijn bij de kwaliteitscriteria Herhaalbaarheid en Vergelijkbaarheid ook geen aanwijzingen voor verschillen gevonden ($p < .10$). Bij Eerlijkheid is de score van bgv-docenten in de tijd toegenomen en die van de avo-docenten gedaald ($p=.088$, $F(1,13)=3.41$, $\eta^2=.208$). Het verschil tussen de A- en B-vraag is op dit kwaliteitscriterium voor beide groepen kleiner geworden ($p=.053$, $F(1,13)=4.53$, $\eta^2=.258$).

Het kwaliteitscriterium Transparantie wordt in de posttest door de avo-docenten minder belangrijk gevonden dan in de pretest, terwijl de bgv-docenten dit duidelijk veel belangrijker zijn gaan vinden (A-vraag; $p=.003$, $F(1,13)=12.85$, $\eta^2=.497$). Uit de analyse van de B-vraag blijkt dat de docenten uit beide groepen wel vinden dat ze in de loop van de tijd meer transparantie hebben gerealiseerd ($p=.018$, $F(1,13)=7.29$, $\eta^2=.359$), waardoor er een trend is dat het verschil tussen gewenst en gerealiseerd (A-vraag minus B-vraag) bij avo sterk is afgenomen en bij bgv min of meer gelijk is gebleven ($p=.089$, $F(1,13)=3.38$, $\eta^2=.206$).

De open vragen uit de posttest zijn door alle negen personen uit de experimentele groep beantwoord.

Zeven van hen (77,8%; 3 bgv- en 4 avo-docenten) gaven aan in de voorafgaande periode ervaringen te hebben opgedaan met het beoordelen van competenties van leerlingen met behulp van rubrics. Alle respondenten zagen mogelijkheden in het gebruik van dergelijke competentiebeoordelingen in hun eigen lessituaties. Op de vraag “welke mogelijkheden zie je” werden de volgende punten genoemd: samenwerking met andere vakken; samenwerken / elkaar helpen; de vorderingen per leerling volgen; meer zelfstandig werken; goede mogelijkheden bij (grote) praktische opdrachten en buitenlessen; en te gebruiken bij het presenteren of een minilesje doen (± 15 min) in een eigen onderwerp van de leerling.

Als beperking is vijf keer de factor tijd genoemd. Daarnaast is aangegeven dat het lastig is om individuele startniveaus van leerlingen vast te stellen. Ook het geven van een theoretisch vak werd als beperking gezien.

Op de (slot-)vraag of competentiebeoordelingen met behulp van rubrics op 't Beeckland structureel toegepast zouden kunnen worden, werd zeven keer met ‘ja’ beantwoord (3 avo- en 4 bgv-docenten), één keer met ‘nee’ (avo-docent), en één keer met ‘weet niet’ (avo-docent).

Conclusie en discussie

Nadat docenten ervaringen hebben opgedaan met het beoordelen van leerlingen met behulp van rubrics, heeft het overgrote deel positief geantwoord op de vraag of dergelijke competentiebeoordelingen te gebruiken zijn op 't Beeckland. Hierbij valt het op dat alle bgv-docenten volmondig 'ja' zeggen en dat er bij de avo-docenten enige weerstand en twijfel is.

Uit de inventarisatie bij de pretest-enquête blijkt dat een groot deel van het personeel van 't Beeckland al iets aan competentiebeoordelingen van leerlingen doet in hun eigen lessituatie. Van de daarbij genoemde voorbeelden is de competentie 'toepassen van kennis' niet genoemd door bgv-docenten, maar wel door avo-docenten. De avo-docenten geven daarmee aan dat zij vinden dat het bij hun vakken vooral om kennis gaat. "Biologie is een kennisvak!", schreef één van hen bovenaan de ingevulde enquête. De leerlingen zullen bij die vakken dan ook voornamelijk aan de hand van kennistoetsen beoordeeld worden. Die leveren min of meer 'harde' gegevens op over geïsoleerde aspecten van kennis. Er zijn wel voorbeelden te vinden waar vaardigheden in het avo worden beoordeeld, bijvoorbeeld luistertoetsen bij het talenonderwijs, maar veel avo-docenten hebben daar geen of weinig ervaring mee (Kerkhoffs, Stark, & Zeelenberg, 2006).

De visie op het leren op 't Beeckland is tot nu toe vooral beperkt geweest tot het opdoen van kennis. In de schoolgids staat dan ook: "Het verwerven van kennis is voor leerlingen de belangrijkste reden om de school te bezoeken. [...] Leerlingen spelen een actieve rol in het proces van kennisverwerving: zij nemen niet alleen leerstof tot zich, maar zij ontwikkelen bovendien de vaardigheden om zelfstandig kennis te verwerven." (Ulenhofcollege, 2011). In de laatste jaren is deze focus geleidelijk aan het verschuiven naar steeds meer competentiegericht onderwijs. In het nog in ontwikkeling zijnde nieuwe schoolontwikkelingsplan is er dan ook structureel meer aandacht voor het ontwikkelen van vaardigheden anders dan het opdoen van kennis.

De invoering van meer competentiegericht onderwijs is bewust gekoppeld aan de geplande invoering van een intersectoraal programma. In dit nieuw te ontwikkelen programma wordt nadrukkelijk gewerkt met de competentieontwikkeling van leerlingen, en daar hoort een nieuwe vorm van beoordelen bij. Het onderwijsgevend personeel dient op deze ontwikkelingen voorbereid te zijn. In gesprekken met docenten wordt aangegeven dat men onbekend is met het beoordelen van competenties. Maar men geeft aan dat er wel behoefte is aan een heldere, eenvoudige, gefundeerde beoordelingswijze. Want veruit de meeste docenten vallen nu weer terug op de vertrouwde (cognitieve) toetsen zodra er beoordeeld moet worden. Mogelijk dat daarbij ook de angst voor het onbekende, een nieuwe manier van werken, een rol speelt.

Er is door de meeste leden van de experimentele groep daadwerkelijk ervaring opgedaan in een eigen lessituatie met een competentiebeoordeling met de rubrics. Maar ook hier blijkt dat sommige docenten alleen aan kennisverwerving denken (inclusief bijbehorende toetsen en schriftelijk beoordelen) als ze aangeven dergelijke competentiebeoordelingen te doen.

Ook wordt er op 'gevoel' gewerkt: zo is er een docent die opschrijft dat er een punt aftrek volgt bij verkeerd gedrag. Zo'n reactie sluit aan bij de waarneming van Korthagen (2004), die aangeeft dat er vaak met de mond wordt beleden dat er competentiegericht wordt opgeleid, maar dat de beoordeling dikwijls gedragsgericht en smal is. Ook zal meespelen dat in de gedachten en meningen over 'goed onderwijs' in het voortgezet onderwijs nog steeds het idee overheerst dat er sprake is van goed onderwijs wanneer door de opleiding een kwantitatief oordeel over de leerling is gegeven. Dit is gebaseerd op een summatief curriculum waarin alles gecontroleerd en beoordeeld wordt. Dit verdraagt zich slecht met een formatief en open leerklimaat dat noodzakelijk is voor de ontwikkeling van competenties (Haacke & van der Ploeg, 2004).

Tijdens de workshop zijn onder andere rubrics van de competentie 'samenwerken' (afbeelding 1) uitgedeeld om mee te oefenen. Bij het uitdelen werd door één van de docenten gezegd: "Dat is handig: daar kan ik wel wat mee!" Deze reactie, die ook door andere aanwezigen werd ondersteund, komt tegemoet aan de wens die Andrade in 1997 al heeft geuit: "Students, as well as teachers, should respond to the use of rubrics by thinking, 'Yes, this is what I need!'" (Andrade, 1997).

Al tijdens de workshop, maar ook in de weken daarna tijdens het gebruik in de eigen lessituatie, zijn er bij de docenten vragen gerezen over de gebruikte Groene Werkplek/CPS-rubrics. Zo is geconstateerd dat deze rubrics heel erg talig zijn. Zeker voor het gebruik met en door vmbo-leerlingen blijkt dit nadelig te zijn: veel leerlingen hebben moeite om de inhoud van de verschillende verwervingsstappen te begrijpen. Daarom is het idee geopperd om te onderzoeken of er andere weergave-vormen ontwikkeld zijn, en zo ja, of die beschikbaar en wellicht beter te gebruiken zijn. Zo is er bijvoorbeeld het idee en initiatief om de competentiebeschrijvingen en de bijbehorende verwervingsstappen te visualiseren in plaatjes (Barnier, 2011).

Een andere vraag en discussiepunt gaat over het streefniveau per leerweg. Bij de gebruikte rubrics wordt uitgegaan van drie leerwegen: BBL, KBL en GTL (afbeelding 1). Hier lijkt een waardeoordeel in te zitten. Een leerling is immers niet bij voorbaat te categoriseren. Door deze CPS-rubrics te gebruiken zal een BBL leerling en zijn docent vooral of alleen maar naar de rubrics kijken behorende bij de BBL-schaal. Is het niet beter om er één rubrics-lijst van te maken, aangezien er toch al overlap in zit? Dan kan een leerling ook groei (blijven) zien als hij op het streefniveau zit van de betreffende leerweg zoals dat nu beschreven is.

Ook is geconstateerd dat competentiebeoordelingen zonder enige vorm van reflectie geen waarde hebben. Het voeren van competentiegesprekken met leerlingen moet daarom structureel aandacht krijgen. De docenten vragen zich af hoe en wanneer de reflectiegesprekken georganiseerd moeten worden. Bij de invoering van een intersectoraal programma zal dit onderwerp speciale aandacht moeten krijgen.

Er zijn verschillende redenen om de resultaten van de analyses van de gesloten vragen voorzichtig te interpreteren. Onder slechte omstandigheden kan het voorkomen dat er geconcludeerd wordt dat de interventie niet werkte terwijl het in feite wel deed, of omgekeerd.

Zo is een Non Equivalent Groups Design gevoelig voor (gebrek aan) interne validiteit: verschillen tussen groepen die al voor het experiment bestonden kunnen de uitkomst beïnvloeden (Trochim, 2006). Dat de groepen al voor het experiment verschillend waren, is hier zeker aan de orde. De experimentele groep bestaat uit docenten die al bewust (in meer of mindere mate) met competentiebeoordelingen bezig waren voor het experiment begon. Zij zijn afkomstig uit een team waarin al in 2010 de keuze is gemaakt om met het onderwerp competentiebeoordelingen aan de slag te gaan, en/of uit een onderwijsontwikkelgroep die werkt aan de invoering van een intersectoraal programma. Een uitgangspunt daarbij is dat competentiegerichte werken een nadrukkelijk onderdeel van is van zo'n intersectoraal programma, en dus ook het competentiegerichte beoordelen.

In zowel de controlegroep als de experimentele groep is het aantal avo-docenten ongeveer even groot als het aantal bgv-docenten. Maar in de gehele schoolpopulatie ligt de verhouding avo- en bgv-docenten op ongeveer 3:1. In het experiment zijn de bgv-docenten derhalve verhoudingsgewijs oververtegenwoordigd.

Uit de analyses van de gesloten vragen blijkt weliswaar dat de absolute scores op de stellingen met betrekking tot belangrijkheid (de A-vragen) bij de kwaliteitscriteria Eerlijkheid, Herhaalbaarheid en Vergelijkbaarheid in de posttest zijn toegenomen ten opzichte van de pretest, maar alleen bij Vergelijkbaarheid is deze trend in de analyse enigszins betrouwbaar gebleken. Ook bij de B-vragen (in hoeverre vind je dat je het onderwerp in de stelling hebt gerealiseerd) zijn de absolute scores overal toegenomen, en ook hier is geen significantie gevonden. Door de toenames lijkt het er op dat er in de tijd wel iets is veranderd, want er is hoger gescoord. In ieder geval is er geen betrouwbaar verschil gevonden tussen de controlegroep en de experimentele groep en kan er worden geconcludeerd dat er geen effect van de interventie is waargenomen. Deels zou dit verklaard kunnen worden uit de korte looptijd van het experiment: met drie maanden was deze erg kort.

Ook speelt mee dat het eventuele effect van de interventie mogelijk vertroebeld is, omdat er tijdens uitvoering interactie (uitwisseling) op werkvloer heeft plaatsgevonden. In de lessen van de beroepsgerichte vakken werken de docenten nauw samen. Men ziet dus zeer goed waar de collega's mee bezig zijn en als er wordt geëxperimenteerd met beoordelingen met rubrics, kan dat de collega's uit de controlegroep niet zijn ontgaan. Avo-docenten werken daarentegen veel meer individueel. Maar ook zij hebben door het uitwisselen van de ervaringen in onder andere de teamvergaderingen toch iets meegekregen: iets uitproberen op een relatief kleine school als 't Beeckland zonder opgemerkt te worden is vrijwel onmogelijk. Daarom is de vraag gerechtvaardigd of de controlegroep in dit experiment wel een echte controlegroep is geweest.

Bij de kwaliteitscriteria Herhaalbaarheid en Vergelijkbaarheid uit de gesloten vragenlijst zijn er geen verschillen gevonden in de scores van de avo- en bgv-docenten.

Bij Eerlijkheid is de score van bgv-docenten in de tijd toegenomen en die van de avo-docenten gedaald. Blijkbaar zijn de bgv-docenten het belangrijker gaan vinden dat het niveau aansluit bij de individuele leerling, en dat er een objectieve beoordeling bij hoort van onderdelen waarvan van tevoren is aangegeven welke belangrijk zijn. Bij beide groepen docenten is het verschil tussen belangrijkheid en gerealiseerd (het verschil tussen de A- en de B-vraag) kleiner geworden.

Bij het kwaliteitscriterium Transparantie wordt gevraagd of leerlingen vooraf weten of de beoordeling meetelt en wat dan de criteria zijn en de procedure is. Dit kwaliteitscriterium wordt door de bgv-docenten in de posttest duidelijk belangrijker gevonden dan in de pretest, terwijl de avo-docenten dit minder belangrijk zijn gaan vinden. Er is een trend dat het verschil tussen de A- en B-vraag bij avo duidelijk is afgenomen, terwijl dat verschil bij de bgv-docenten min of meer gelijk is gebleven.

Als gevraagd wordt naar aanbevelingen voor het vervolgtraject voor invoering van competentiebeoordelingen op 't Beeckland, wordt door veel docenten de factor tijd genoemd. Blijkbaar wordt dit als wezenlijk knelpunt ervaren. Op zich is dat niet uniek: ook op andere scholen is geconstateerd dat rubrics tijd

kosten, zowel tijdens de voorbereiding als bij de uitvoering en dat die extra tijdsinvestering vooral neerkomt op de docenten (Kerkhoffs, Stark, & Zeelenberg, 2006). Een deel van dit probleem kan door het schoolmanagement worden ondervangen, omdat zij daarin een faciliterende taak heeft. Kerkhoffs et al. (2006) noemen daarvoor een aantal mogelijkheden, zoals docenten vrijstellen van lestaken, roosteraanpassingen, bij- en nascholing of deskundige begeleiding inkopen. Twee andere mogelijkheden die door de docenten uit de experimentele groep zijn genoemd zijn het integreren van vakken en het verdelen van de competenties over vakken.

Andere aanbevelingen die genoemd zijn: leg de beoordelingscriteria vast, zodat alle leerlingen schoolbreed gelijk worden beoordeeld (twee keer genoemd), en creëer draagvlak, bijvoorbeeld door begeleiding van de docenten en (het organiseren) van collegiale intervisie.

Hoewel de traditionele vorm van beoordelen voor veel docenten nog steeds erg belangrijk is en blijft, reageren vanuit de meeste docenten uit de experimentele groep positief op het voorstel om competentiebeoordelingen voor leerlingen in te passen in een nieuw te ontwikkelen intersectoraal programma. Als er bij dat proces structureel aandacht is voor een aantal essentiële randvoorwaarden, zoals begeleiding, scholing en tijd, lijkt er dan ook voldoende draagvlak te zijn om er daadwerkelijk mee te starten.

Referenties

- Andrade, G. H. (1997). Understanding Rubrics. *Educational Leadership*, 54, 4. Ontleend aan <http://learnweb.harvard.edu/alps/thinking/docs/rubricar.htm>
- Baarda, D.B., Goede, M.P.M. de, & Dijkum, C.J. van (2007). *Basisboek statistiek met SPSS. Handleiding voor het verwerken en analyseren en rapporteren over (onderzoeks)gegevens*. Houten: Wolters-Noordhoff.
- Baartman, L., Bastiaens, T., Kirschner, P., & Geurts, L. (2005). *Resultaten van het onderzoek naar kwaliteitscriteria voor Competentie Assessment Programma's*. Heerlen: Open Universiteit Nederland.
- Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & Vleuten, C. P. M. van der (2006). The wheel of competency assessment: presenting quality criteria for Competency Assessment Programs. *Studies in Educational Evaluation*, 32, 153-177.
- Baartman, L., Seezink, A., Sanden, J. van der, & Brandt, C. van der (2006). *De VMBO ICT-route. Resultaten van een onderzoek naar het vormgeven van competentiegericht toetsing en krachtige leeromgevingen*. Heerlen/Eindhoven/Tilburg: Open Universiteit Nederland, Technische Universiteit Eindhoven, Universiteit van Tilburg.
- Barnier, L. (2011). Zo ziet competent gedrag eruit. Plaatjes helpen docenten bij coachingsgesprek. *Vakblad Groen Onderwijs* 14, 22-24.
- Berg, R. van den, & Vernooij, K. (2000). Implementatie van onderwijsinnovaties: naar een samenhangende benadering. In: Creemers, B.P.M. & Houtveen, A.A.M. (2000). *Onderwijsinnovatie, onderwijskundig Lexicon*. Alphen aan den Rijn: Kluwer.
- Bergen, T., & Veen, K. van (2004). Het leren van leraren in een context van onderwijsvernieuwingen: waarom is het zo moeilijk? *VELON Tijdschrift voor lerarenopleiders*, 25(4), 29-39.
- Bijsterveldt, M. van (2011). Op weg naar een toekomstbestendig vmbo. *Beleidsbrief 29 april 2011*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- Boeije, H. (2005). *Analyseren in kwalitatief onderzoek*. Amsterdam: Boom Onderwijs.
- Boer, E. de, & Kuhlemeier, H. (2005). *Competentiegericht leren en beoordelen in vmbo en mbo*. Utrecht: WVOI.
- Boer, P. den, Mittendorff, K., & Sjenitzer, T. (2004). *Beter Kiezen. Onderzoek naar keuzeprocessen van jongeren in het VMBO en MBO*. Wageningen: Stoas Onderzoek.
- Bolt, L. van der, Studulski, F., Vegt, A.L. van der, & Bontje, D. (2006). *De betrokkenheid van de leraar bij onderwijsinnovaties. Een verkenning op basis van literatuur*. Utrecht: Sardes.
- CBS (2011). *Jaarboek onderwijs in cijfers 2011*. Den Haag/Heerlen: Centraal Bureau voor de Statistiek.
- Cooper, B.S., & Gargan, A. (2009). Rubrics in education. Old term, new meanings. *Phi Delta Kappan*, 91(1), 54-55.
- Dijkstra, A., & Baartman, L. (2011). Zelfevaluatie van de kwaliteit van assessment. *OnderwijsInnovatie*, maart 2011, 17-25.
- Dochy, F., & Nickmans, G. (2005). *Competentiegericht opleiden en toetsen. Theorie en praktijk van flexibel leren*. Utrecht: Lemma.
- Field, A. (2009). *Discovering statistics using SPSS*. Londen: Sage.
- Gulikers, J. (2008). *Assessment in doorlopende leerlijnen*. [PowerPoint-presentatie] Wageningen: 4C Your Way Conferentie, 16 December 2008. Ontleend aan <http://www.4cyourway.nl/Uploadsvanderden/tabid/1258/Default.aspx>
- Haacke, R., & Ploeg, Y. van der (2004). Reflectie en portfolio en reflectie, wat te beoordelen? In: *Eindrapport Kenniskring Reflectie op het handelen*, juni 2004, 38-46. Amsterdam: Hogeschool van Amsterdam.

- Hamstra, D.G., & Ende, J. van den (2006). *De vmbo-leerling. Onderwijspedagogische- en ontwikkelingspsychologische theorieën*. Amersfoort: CPS.
- Illeris, K. (2002). *Three dimensions of learning*. Roskilde, Denmark: Roskilde University Press / Leicester, UK: NIACE.
- Jong, C. de (2008). *Overzicht verwervingsstappen competenties BB en KB vmbo Groen*. Ontleend aan <http://www.4cyourway.nl/Uploadsvanderden/tabid/1258/Default.aspx>
- Jong, C. de, Hoekstra, T., Wel, M. van der, & Wensink, J. (2005). *Competentieleren in het vmbo. Ontwikkelingsvoorbeelden uit de praktijk*. Amersfoort: CPS.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130-144.
- Kerkhoffs, J., Stark, E., & Zeelenberg, J. (2006). *Rubrics als beoordelingsinstrument voor vaardigheden*. Enschede: SLO.
- Kirschner, P.A. (2006). *(Inter)dependent learning: Learning is interaction*. (Inauguratierede). Utrecht University, Utrecht.
- Korthagen, F. (2004). Zin en onzin van competentiegericht opleiden. *VELON Tijdschrift voor Lerarenopleiders*, 25(1), 13-23.
- Nijman, J. (2008). De groene werkplek. Competentiegericht leren in het vmbo als project. *Vakblad Groen Onderwijs*, 7, 10-12.
- Onderwijsraad (2009). *Examens in het vmbo*. Advies Nr. 20090116/960, uitgebracht aan de staatssecretaris van Onderwijs, Cultuur en Wetenschap. Den Haag: Onderwijsraad.
- Regeling intra- en intersectorale programma's v.m.b.o. (2008). *Staatscourant* 2008, nr. 140.
- Seligman, M.E.P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist* 55(1), 5-14.
- Sinke, G.P.J. (2006). *Aan de slag met assessment. Toetsen en beoordelen in een competentiegerichte leeromgeving*. Nuenen: Onderwijsadviesbureau Dekkers.
- Sluijsmans, D. (2008). *Betrokken bij beoordelen. Intreerede lectoraat Duurzaam beoordelen in vraaggestuurd leren*. Nijmegen: Hogeschool van Arnhem en Nijmegen.
- Steeg, M. van der, Vermeer, N., & Lanser, D. (2011). Nederlandse onderwijsprestaties in perspectief. *CPB Policy Brief* 2011/05.
- Trochim, W.M.K. (2006). The Nonequivalent Groups Design. *Research Methods Knowledge Base*. Ontleend aan <http://www.socialresearchmethods.net/kb/quasnegd.php>
- Ulenhofcollege (2011). *Schoolgids*. Doetinchem: Ulenhofcollege. Ontleend aan <http://www.ulenhof.nl/Algemeen/SchoolgidsInhoud/SchoolgidsAlgemenegegevens/tabid/777/Default.aspx>
- Valcke, M. (2010). *Onderwijskunde als ontwerpwetenschap. Een inleiding voor ontwikkelaars van instructie en voor toekomstige leraren*. Gent: Academia Press.
- Ward, M., Gruppen, L., & Regehr, G. (2002). Measuring self-assessment: Current state of the art. *Advances in Health Science Education*, 7, 63-80.
- Waslander, S. (2007). *Leren over innoveren. Overzichtsstudie van wetenschappelijk onderzoek naar duurzaam vernieuwen in het voortgezet onderwijs*. Utrecht: VO-project Innovatie. Ontleend aan <http://www.vo-raad.nl/publicaties/brochures/leren-over-innoveren>
- Wells, C.S., & Wollack, J.A. (2003). *An Instructor's Guide to Understanding Test Reliability*. Madison: University of Wisconsin. Ontleend aan <http://testing.wisc.edu/Reliability.pdf>
- Wester, M., & Kessel, N. van (2012). *In- en doorstroom sectoroverstijgende programma's in het vmbo. Cohort 2008 en 2009*. Nijmegen: ITS Radboud Universiteit.
- Wikeley, F., Stoll, L., Murillo, J., & Jong, R. de (2005). Evaluating effective school improvement: Case studies of programmes in eight European countries and their contribution to the Effective School Improvement Model. *School Effectiveness and School improvement*, 16, 387-405.
- Wout, E. van 't, & Jong, C. de (2010). *De Groene Werkplek: praktisch materiaal voor competentieleren*. [PowerPoint-presentatie] Bunschoten: Groen Proeven 'Doe dag', 18 maart 2010. Ontleend aan <http://ebookpp.com/pl/plotterweg-ppt.html>

Dankwoord

Zonder de hulp en medewerking van velen zou de voltooiing van deze thesis niet gelukt zijn.

Zoals het thuisfront (Dinie, Miriam en Michiel), de collega's van 't Beeckland, de Stoas-docenten, de mede-masterstudenten en alle andere belangstellenden.

Sommigen verdienen een speciale vermelding:

Sjaak Stans, voor zijn rol als opdrachtgever, zijn ondersteuning en voor het beschikbaar stellen van tijd (een lesvrije vrijdag): zonder deze wekelijkse studiedag zou het mij niet gelukt zijn deze studie binnen de daarvoor gestelde tijd af te krijgen.

Judith Gulikers, voor haar waardevolle tips bij de 'go/no go' voor de start van het onderzoek.

De leerlingen van B4B en K4D (afdeling Groen), voor de bereidwillendheid mee te werken aan de video-opnames en voor het geven van duidelijke feedback op het gebruik van rubrics.

Marcel Overvelde, voor het maken van de video-opnames.

Martijn van Huijgevoort, voor de ondersteuning bij de statistische analyses.

Madelon Bontje, voor het aanleveren van artikelen en voor haar feedback.

Jorien Nijhuis, voor het verbeteren van de Engelse samenvatting.

Gentle Giant, Phideaux, IOEarth, Spinvis, Tangerine Dream, Ayreon, e.v.a., voor de inspiratie.

En Loes Spit natuurlijk. Voor haar inbreng, begeleiding en ondersteuning kan ik met twee woorden volstaan: *"Helemaal goud"*.

Bedankt allemaal!

BIJLAGE 1. Vragenlijst pretest

A) Open vragen

Beoordeel je in jouw lessen competenties van leerlingen? Ja/nee

Zo ja:

- Welke competenties worden beoordeeld?
- Op welke manier vinden deze beoordelingen plaats?
- Welk materiaal gebruik je voor de beoordeling? (scorelijsten, lesmethode e.d.)

Zo nee:

- Zou je competenties van leerlingen willen beoordelen? Waarom wel/niet?

B) Gesloten vragen

Voor de volgende stellingen neem je een opdracht uit één van jouw lessen in gedachte waarbij je een competentie beoordeelt. Bijvoorbeeld als leerlingen samenwerken, iets presenteren of een beroepshandeling uitvoeren.

Korte beschrijving van deze opdracht:

Omcirkel de best passende score.

1 = niet, 2 = nauwelijks, 3 = neutraal, 4 = enigszins, 5 = wel

1 De beoordelingsopdracht sluit aan bij het niveau van de individuele leerling.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

2 Er is vooraf bepaald welke onderdelen belangrijk zijn en zwaarder meetellen in de eindbeoordeling.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

3 Subjectieve of niet in de criteria opgenomen zaken tellen niet mee voor de beoordeling.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

4 Leerlingen weten vooraf of de beoordeling meetelt voor het rapport/schoolexamen.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

5 Voor leerlingen is vooraf duidelijk wat de beoordelingscriteria zijn.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

Competentiebeoordelingen met rubrics op 't Beeckland

6 Voor leerlingen is het duidelijk wat de beoordelingsprocedure is.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

7 De beoordelingscriteria zijn gericht op direct waarneembare zaken.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

8 Door de manier van beoordelen blijkt direct hoe een leerling zal functioneren in een beroep.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

9 Als twee personen de opdracht beoordelen (tegelijkertijd en/of in verschillende klassen), wordt gecontroleerd of hun beoordelingen overeenkomen.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

10 Als de beoordeling herhaald wordt, wordt gecontroleerd of dit dezelfde uitkomst oplevert.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

11 De beoordelingsopdrachten zijn voor alle leerlingen gelijk.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

12 De beoordelingscriteria zijn voor alle leerlingen gelijk.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

13 Alle leerlingen worden in een vergelijkbare situatie beoordeeld.

Hoe belangrijk vind je dit bij de competentiebeoordeling? 1 2 3 4 5

In hoeverre vind je dat je dit hebt gerealiseerd? 1 2 3 4 5

BIJLAGE 2. Vragenlijst posttest

A) Open vragen

Wat zijn jouw ervaringen met het beoordelen van competenties van leerlingen?

Welke mogelijkheden zie je?

Welke beperkingen zie je?

We hebben het gehad over 'rubrics', een soort scorelijst voor competenties. Als voorbeeld zijn CPS-rubrics uitgedeeld. Hebben deze rubrics voor jou een toegevoegde waarde gehad? Zo ja: welke? Zo nee: waarom niet?

Als we standaard aandacht willen geven aan de beoordeling van competenties van leerlingen, waar moeten we dan vooral op letten?

Kunnen of moeten we op 't Beeckland dergelijke 'rubrics' structureel toepassen?

B) Gesloten vragen

De gesloten vragen van de posttest zijn identiek aan die van de pretest.