

De beoordeling van startbekwaamheden van beginnende eerstegraads docenten Engels

ANNE MARIE UHLENBECK

De ontwikkeling van een beroepsstandaard voor docenten en de beschrijving van startbekwaamheden roept de vraag op hoe na te gaan of docenten beschikken over de beschreven bekwaamheden. Dit artikel beschrijft de eerste aanzet tot een beoordelingsprocedure voor beginnende docenten Engels: wat wordt er beoordeeld, met welke instrumenten en welke personen zijn betrokken bij de ontwikkeling van zo'n beoordelingsprocedure.

In de toekomst krijgen beginnende docenten hun bevoegdheid op grond van aantoonbare startbekwaamheid en niet meer op grond van het succesvol doorlopen van de lerarenopleiding. Deze startbekwaamheden zijn beschreven (VSLPC 1999) en de vraag is vervolgens hoe vast te stellen of iemand startbekwaam is. Mijn onderzoek betreft de ontwikkeling van een beoordelingsprocedure toegespitst op beginnende eerstegraads docenten Engels. De keuze om instrumenten te ontwikkelen voor één vak en voor één niveau is gebaseerd op twee vooronderstellingen. Startbekwaamheden blijken samen te hangen met de leeftijd van de leerlingen en met de inhoud van het vak: een docent Engels in het havo/vwo heeft andere kennis en vaardigheden dan een docent wiskunde in het vmbo (Shulman

1987). In de tweede plaats staat de ontwikkeling van een beoordelingsprocedure voor één vak en één niveau model voor de ontwikkeling van procedures voor andere vakken en niveaus omdat de problemen bij de ontwikkeling van beoordelingsinstrumenten niet fundamenteel verschillen (Verloop 1994). Het uiteindelijke doel van de beoordeling is om een beslissing te kunnen nemen of de beginnende docent startbekwaam is of nog niet. Tevens moet de beoordeling een bijdrage leveren aan de professionalisering van de docent door aanknopingspunten te bieden voor verdere ontwikkeling. Of de beoordeling het beste aan het einde van de opleiding zou kunnen plaatsvinden of als de docent al is begonnen met werken, is nog niet duidelijk.

Uitgangspunten

Als eerste stap in de ontwikkeling van de beoordelingsprocedure heb ik een raamwerk van eisen opgesteld dat aangeeft waaraan een beoordelingsprocedure voor (beginnende) docenten moet voldoen. De in totaal 15 eisen zijn ontleend aan huidige inzichten met betrekking tot de aard van het onderwijzen, hoe docenten leren en zich ontwikkelen en wat docenten moeten weten om goed les te geven (Borko & Putnam, 1996; Calderhead,

1996; Doyle, 1990; Freeman & Richards, 1996; Woods, 1996). Daarnaast is recent onderzoek met betrekking tot het beoordelen van beroepsvaardigheden geraadpleegd (Andrew & Barnes 1990; Dwyer & Stufflebeam 1996; Haertel, 1991). Het raamwerk vormt het referentiekader in het onderzoek.

Hieronder bespreek ik ter illustratie een deel van het raamwerk en formuleer ik vervolgens de implicaties daarvan voor de beoordeling van beginnende docenten.

- *Lesgeven is een complexe vaardigheid.* Lesgeven heeft de intentie veranderingen te bewerkstelligen in het denken en handelen van groepen leerlingen die onderling verschillen wat betreft hun individuele kenmerken, hun behoeften en achtergronden. Lesgeven vindt bovendien plaats in een dynamische omgeving die voor een deel onvoorspelbaar is en daardoor onzekerheid met zich meebrengt en waarin veel gelijktijdig plaats vindt waarop onmiddellijk moet worden gereageerd. Lesgeven vraagt van de docent een goed inzicht in de situatie en het beschikbaar hebben van een repertoire aan handelingsalternatieven (Doyle 1990).
- *Lesgeven vindt altijd plaats in een context.* Deze context - de specifieke lesinhoud, het type leerlingen, de klas en het type school - oefent aanzienlijk invloed uit op de beslissingen van de docent. De docent op zijn of haar beurt oefent invloed uit op de situatie. Lesgeven kan niet los worden gezien van de context. Wat effectief en passend onderwijsgedrag is, wordt bepaald door de situatie en de doelen die de docent nastreeft. Wat goed is in de ene context hoeft dat niet te zijn in een andere (Beijaard & Verloop 1996; Freeman & Richards 1996).
- *Lesgeven is niet alleen handelen, maar handelen op basis van inzichten.* De docent handelt op basis van kennis van de vakinhoud, wat leerlingen boeit, wat zij moeilijk vinden, hoe lang zij nodig hebben om bepaalde taken uit te voeren en op basis van inzichten hoe bepaalde

lesinhoud het beste aan bod kan komen. Docenten kunnen dezelfde doelen nastreven maar de strategieën die zij hanteren, kunnen verschillen, afhankelijk van hun persoonlijke opvattingen over leren en onderwijzen en van hun interpretatie van de situatie (Calderhead 1996; Verloop, Beijaard, & Driel 1998; Woods 1996).

- *Lesgeven bestaat uit meer dan alleen het uitvoeren van activiteiten in de klas.* Ook voorbereidende activiteiten en evaluatieve activiteiten na afloop van een les of een serie lessen (alleen of in samenwerking met collega's) maken deel uit van lesgeven (Dwyer & Stufflebeam 1996).
- *Er zijn verschillende soorten kennis waarover een docent moet beschikken om goed les te kunnen geven.* De docent moet gebruik kunnen maken van inzichten afkomstig uit onderzoek, zoals kennis over het leren en onderwijzen van vreemde talen. Vanuit dit perspectief moeten beginnende docenten laten zien dat ze bijvoorbeeld bepaalde onderwijsstrategieën kunnen toepassen en aanpassen aan hun eigen situatie. Criteria ontleend aan onderzoeksliteratuur worden gebruikt om te beoordelen hoe goed een beginnend docent dit doet. Vanuit een praktijkgericht perspectief moet de docent praktijkkennis ontwikkelen. Praktijkkennis verwijst naar kennis van klassensituaties en praktische dilemma's van het lesgeven. Ervaringen opgedaan in de klas en vooral reflectie op deze ervaringen vormen de voornaamste bron voor de ontwikkeling van deze kennis. Criteria waarop de beginnende docent wordt beoordeeld, worden ontleend aan wat ervaren docenten goed vinden (Cochran-Smith & Lytle 1999; Meijer 1999).

Aan bovengenoemde inzichten ontleen ik vijf belangrijke implicaties voor de beoordeling van beginnende docenten:

- de beoordeling richt zich op omvattende taakgebieden die de complexiteit van het

lesgeven benaderen en niet op geïsoleerde vaardigheden;

- de beoordeling vindt zoveel mogelijk plaats in contexten die lijken op de werkelijkheid;
- de beoordeling richt zich op een combinatie van observeerbaar gedrag en achterliggende kennis en opvattingen;
- een brede opvatting over lesgeven houdt de ontwikkeling van meerdere instrumenten in;
- de criteria waarop beginnende docenten worden beoordeeld, zijn vastgesteld in een dialoog tussen inzichten uit onderzoek en opvattingen van ervaren docenten.

Wijze van aanpak

De ontwikkeling van de beoordelingsinstrumenten en de overwegingen die daarbij een rol hebben gespeeld, vormen de hoofdmoot van het onderzoek.

Ontwikkelteam en assessoren

Vijf ervaren docenten Engels, waaronder één lerarenopleider waren gedurende 18 maanden één dag in de week betrokken bij het onderzoek. Zij vormden het ontwikkelteam en traden op als beoordelaar bij het testen van de verschillende instrumenten. In een later stadium zijn nog eens twee docenten Engels en een lerarenopleider betrokken bij het onderzoek. Zij traden op als beoordelaar na eerst een training gevolgd te hebben.

Vaststellen van startbekwaamheid van beginnende docenten Engels

Voor een nadere omschrijving van de startbekwaamheden van docenten Engels heb ik documenten zoals het Beroepsprofiel, Raamplan voor de Universitaire Lerarenopleiding, de kerndoelen Engels in de Basisvorming en eindexamen programma's voor het vak Engels geraadpleegd. Daarnaast heb ik gebruik gemaakt van onderzoeksliteratuur over wat bekend is over goed onderwijzen, zoals studies

over inhoud van de kennis van leraren en studies op het gebied van het vreemde talenonderwijs. Op basis van een analyse van deze bronnen en op basis van commentaar van het ontwikkelteam zijn startbekwaamheden geformuleerd. Dit proces resulteerde uiteindelijk in de beschrijving van drie domeinen (inclusief subdomeinen) van bekwaamheid:

- domein I, creëren van voorwaarden voor het leren van leerlingen;
- domein II, stimuleren van het leren van leerlingen van de vreemde taal;
- domein III, bevorderen van de eigen professionele groei en verantwoording naar derden.

Beoordelingsinstrumenten

Verschillende overwegingen hebben een rol gespeeld bij de keuze van instrumenten. Elk van de gekozen instrumenten heeft positieve en negatieve eigenschappen ten aanzien van validiteit en betrouwbaarheid. Door instrumenten te combineren heb ik geprobeerd de beperkingen van elk instrument op zich te compenseren. De simulaties en de toets gesloten vragen brengen kandidaten in vergelijkbare situaties. Zij voeren dezelfde taken/opdrachten uit en worden volgens een vaste procedure beoordeeld. De inhoud van de taken/opdrachten zal voor de ene kandidaat echter vertrouwd zijn dan voor de andere, afhankelijk van de ervaring die zij hebben opgedaan. Het portfolio geeft elke kandidaat de gelegenheid om - binnen de gegeven richtlijnen - te bepalen hoe zij hun bekwaamheid demonstreren binnen hun eigen lespraktijk. Hieronder ga ik uitvoeriger in op de instrumenten.

Portfolio. Met behulp van een handleiding documenteren kandidaten hun lespraktijk over een periode van vier tot zes weken. De handleiding geeft aan hoe zij hun lespraktijk kunnen documenteren en geeft aan hoe zij beoordeeld zullen worden. Ze kiezen 5 tot 8 thema's en verzamelen bewijsmaterialen om hun bekwaamheden met betrekking tot de

gekozen thema's te demonstreren. Verplichte bewijsmaterialen zijn: een video van een gegeven les; werk van leerlingen; eigen gemaakt materiaal, zoals lesplannen en opdrachten. De kandidaten lichten in een gesprek de inhoud van hun portfolio nader toe. Het portfolio geeft informatie over wat kandidaten doen in hun eigen onderwijspraktijk met betrekking tot de gekozen thema's. Het geeft ook informatie over de mate waarin zij reflecteren op hun lespraktijk en werken aan hun professionele ontwikkeling. Het portfolio is in twee rondes beproefd. In de eerste ronde bij twee kandidaten; in de tweede ronde bij vijf kandidaten.

Het portfolio kan op een valide manier de complexe lespraktijk in kaart brengen. Het beoordelen van het portfolio is echter problematisch. Zo kan de samenstelling ervan uiteenlopen en kunnen de thema's, de typen bewijsmaterialen en de kwaliteit van de toelichting sterk verschillen. Daarnaast zijn er soms aanzienlijke verschillen in de schoolcontext: sommige kandidaten krijgen goede begeleiding, anderen worden aan hun lot overgelaten. De vraag is dan of het portfolio iets over de kandidaat zegt of over de context waarin de kandidaat opereert.

Drie simulaties van authentieke werksituaties De kandidaat: a) bereidt een les voor met behulp van een leergang voor een brugklas, b) kijkt brieven na van leerlingen van 5 havo en bedenkt een aanpak in de klas en c) werkt samen met een collega in de selectie van lesmateriaal gespreksvaardigheid voor een lessenserie in 4vwo/havo. Alle materialen die onderdeel uitmaken van de simulatie zijn authentiek. In de opdracht is aangegeven op welke criteria de kandidaat wordt beoordeeld. Kandidaten bereiden de opdrachten individueel voor. Na afloop lichten zij in een semi-gestructureerd interview hun keuzes en hun motieven toe. De simulaties geven diepgaande informatie over gebruikskennis van de kandidaten op verschillende subdomeinen. Het gaat hier met name

om de integratie en de uitgebreidheid van deze kennis en inzichten en de mate waarin ze daar gebruik van maken. Bijvoorbeeld de kennis die ze hebben over leerlingen, over hoe je leerlingen kunt motiveren en activeren gekoppeld aan specifieke leerstof, over hun kennis van werkvormen en wanneer en hoe deze te gebruiken. De simulaties zijn in twee rondes beproefd. In de eerste ronde bij dertien kandidaten; in de tweede ronde bij veertien kandidaten.

De simulaties bootsen actuele taken na die leraren uitvoeren als onderdeel van hun werk en zijn zodoende representatief voor lesgeven. Het voordeel van simulaties is dat kandidaten zich in vergelijkbare situaties bevinden en dat de scoringsprocedure een zekere mate van standaardisatie toelaat. Echter, de simulaties staan buiten de eigen schoolcontext en zijn om die reden minder valide. Een ander belangrijk nadeel is dat simulaties tijdrovend zijn om uit te voeren met als gevolg dat slechts een beperkt aantal simulaties kan worden uitgevoerd.

Toets met 40 gesloten vragen. Deze toets stelt uiteenlopende vragen over Domein I en II. Het betreffen vooral kennis- en inzichtvragen. Dit instrument is in twee rondes beproefd. Aan de eerste ronde namen 53 kandidaten deel. Op basis van de ervaringen uit deze eerste ronde zijn de vragen bijgesteld en in een tweede ronde beproefd bij 74 kandidaten. De toets met gesloten vragen is efficiënt, omdat in korte tijd uiteenlopende kennis- en inzichtvragen kunnen worden gesteld. De vragen kunnen betrouwbaar gescoord worden. Het nadeel is dat de relatie met de lespraktijk zwak is.

Training beoordelaars

Het gebruik van de procedure is ingewikkeld. De respons van kandidaten op de verschillende onderdelen is complex en vraagt een grote mate van deskundigheid van de beoordelaars. Een voorwaarde is een grondige ervaring met de lespraktijk en kennis van het vreemde-talen-onderwijs. Om beoordelaars voor te bereiden op hun taak is een training ontwikkeld. De

beoordelaars leren werken met de scoringsprocedure en maken kennis met de instrumenten, de criteria waarop kandidaten worden beoordeeld en de visie op onderwijzen die het uitgangspunt vormt van de beoordeling. Zij moeten leren hun (vaak impliciete) opvattingen over wat zij verstaan onder goed vreemde talenonderwijs te expliciteren. Het doel van de training is te komen tot een gemeenschappelijk *frame-of-reference*, dat wil zeggen een gedeelde visie op (leren) onderwijzen en op de rol van beoordeling daarin.

Feedback

De kandidaten ontvangen een uitgebreid feedbackrapport. Het rapport bestaat uit een beschrijving en commentaar op de wijze waarop de kandidaten op de verschillende onderdelen van de beoordeling hebben gepresteerd. De conclusie bestaat uit een beschrijving van sterke en zwakke punten gekoppeld aan de drie eerder genoemde domeinen.

Beoordelingsprocedure

Tenslotte de beoordelingsprocedure zelf. De kandidaten zijn steeds door twee assessoren beoordeeld. De beoordeling vindt plaats gedurende een periode van ongeveer drie maanden en bevat vijf onderdelen.

- De kandidaten ontvangen schriftelijk instructie over de hele procedure. Zij ontvangen de handleiding voor het samenstellen van het portfolio.
- De kandidaten stellen hun portfolio samen over een periode van vier tot zes weken (tijdsinvestering 10 tot 15 uur).
- De kandidaten komen naar de instelling en leveren hun portfolio in. Zij voeren drie simulaties uit (twee dagdelen).
- De kandidaten komen naar de instelling om hun portfolio nader toe te lichten en de toets gesloten vragen te maken (een dagdeel).
- Enige weken later ontvangen de kandidaten het feedbackrapport.

Resultaten

De drie ontwikkelde instrumenten zijn goed bruikbaar bij de beoordeling van een beginnend docent. Gecombineerd leveren zij 'rijke' en gedetailleerde informatie op over de kandidaat. Met name de simulaties leveren veel informatie op omdat de kandidaten over hun keuzen en beslissingen bevraagd worden. Het portfolio geeft een goed beeld over het functioneren van kandidaten in de school en de thema's die hen bezig houden. Het overheersende beeld van het portfolio is 'werk in uitvoering'. De toets gesloten vragen geeft een beeld van de algemene onderwijskundige en vakdidactische kennis van een kandidaat.

De context waarin de kandidaten zich in de scholen bevinden, verschilt aanzienlijk en dat maakt de beoordeling van met name het portfolio lastig. Een voorbeeld hiervan is een item uit het portfolio van een kandidaat over de steunlessen Engels die zij geeft als onderdeel van haar lesgeeftaak. Vanuit haar sectie of vanuit de school krijgt zij weinig ondersteuning. Zij bedenkt zelf een invulling en experimenteert met veel inzet daarmee. De wijze waarop en de inhoudelijke invulling blijkt echter mager en getuigt niet van veel kennis en inzicht hoe leerlingen het beste te helpen. Hoe dit te beoordelen? Extra moeilijk indien dit voorbeeld wordt afgezet tegen een andere kandidaat die in een sectie en een school werkt die een goed sectiebeleid voert en veel ondersteuning biedt aan beginnende docenten.

De weging van de uitkomsten van de verschillende instrumenten vraagt bijzondere aandacht. Bij het combineren van de informatie uit de drie instrumenten, beschouwen we de drie simulaties, de thema's van het portfolio en de toets gesloten vragen als op zich zelf staande componenten, die elk een ander licht op het functioneren van de kandidaat werpen. Op basis van de oordelen van de assessoren wordt het functioneren van de kandidaat op elke component beschreven. Tenslotte wordt gezocht naar terugkerende patronen over de

verschillende componenten heen en wordt een conclusie in termen van sterktes en zwaktes met betrekking tot het functioneren van de kandidaat op de drie domeinen geformuleerd.

Om na te gaan of de kandidaten zich kunnen herkennen in het feedbackrapport en of het oordeel overeenkomt met hun eigen inschatting van hun functioneren, zijn de kandidaten na ontvangst van het rapport telefonisch geënquêteerd. Over het geheel genomen kunnen de kandidaten zich goed herkennen in de beschrijving van hun sterke en zwakke punten. Zij beoordelen het rapport als informatief en ondersteunend voor hun verdere professionalisering. Ook het doorlopen van de procedure hebben zij als leerzaam ervaren.

De kandidaten vinden de combinatie van instrumenten een zinvolle manier om hun startbekwaamheid te beoordelen. De samenstelling van het portfolio stelt hoge eisen aan de kandidaten: ze moeten geschikte thema's kiezen, bewijsmaterialen zoeken die hun beweringen ondersteunen en tenslotte hun handelen toelichten in de vorm van reflecties. Dit vraagt veel van hun vermogen hun ervaringen schriftelijk onder woorden te brengen.

Bespreking

Het onderzoek naar de ontwikkeling van een beoordelingsprocedure voor beginnende docenten roept vragen op die slechts tentatief te beantwoorden zijn. Kenmerken van goed vreemde talenonderwijs zoals dat uit de literatuur naar voren komt, in de vorm van de 'communicative classroom' waarin leraar en leerlingen en leerlingen onderling actief bezig zijn met de vreemde taal, contrasteert met wat vele docenten vreemde talen in de praktijk doen, met wat zij haalbaar vinden en soms wat zij wenselijk achten. Over de rol van grammatica in het vreemde talenonderwijs, over de mogelijkheden van het aanleren van

strategieën, over de haalbaarheid van het gebruik van de vreemde taal als instructie- en communicatiemiddel enz. bestaan uiteenlopende opvattingen. Dat geldt voor docenten onderling, maar ook in de literatuur bestaan verschillen in opvattingen. Een vraag is hoe met dit gegeven om te gaan en wie of wat de doorslag geeft bij het formuleren van criteria? Een belangrijke ervaring van dit onderzoek is dat concrete beoordelingssituaties een aanleiding vormen om de discussie over goed (vreemde talen) onderwijs aan te gaan.

Hoe van de uitgebreide en gedetailleerde informatie over een kandidaat te komen tot een verantwoorde beslissing - wel of nog niet startbekwaam - is een bijzonder belangrijke vraag. Welke beslisregels worden gehanteerd, uitgaande van het gegeven dat kandidaten een geschakeerd beeld laten zien op de te onderscheiden onderdelen van de beoordeling en dus sterke en zwakkere punten hebben? Indien kandidaten mogen compenseren voor (on)bekwaamheid wordt verondersteld dat zij zelf tekorten kunnen wegwerken. Als daarentegen de eis zou zijn een minimum aan bekwaamheid op verschillende onderdelen, dan bepaalt het slechtste onderdeel of iemand startbekwaam zal worden bevonden. Dit lijkt niet te verdedigen tegenover de kandidaat die op andere onderdelen goed heeft gepresteerd.

Nogmaals, dit onderzoek pretendeert niet een adequaat antwoord op dit soort vragen te hebben. Wel is gepoogd hiermee zo goed mogelijk om te gaan. De uiteindelijke reden voor het ontwikkelen van docentbeoordelingen is de verwachting dat het onderwijs daar mede goede docenten door krijgt. Of deze verwachting juist is, moet nog blijken.

Het onderzoek is medegefinancierd door OC&W. Douwe Beijaard dank ik voor zijn commentaar.

LITERATUUR

- Andrews, T. E., & Barnes, S. (1990). Assessment of teaching. In Houston, W.R. (Ed.), *Handbook of research on teacher education* (pp. 569 - 598). New York: Macmillan.
- Beijaard, D., & Verloop, N. (1996). Assessing teachers' practical knowledge. *Studies in Educational Evaluation*, 22(3), 275-286.
- Borko, H., & Putnam, R. T. (1996). Learning to teach. In Berliner, D.C. & Calfee, D.C. (Eds.) *Handbook of educational psychology* (pp. 673-708). New York: MacMillan.
- Calderhead, J. (1996). Teachers: Beliefs and knowledge. In Berliner, D.C. & Calfee, R.C. (Eds.), *Handbook of educational psychology* (pp. 709-725). New York: MacMillan.
- Cochran-Smith, M., & Lytle, S. L. (1999). Relationships of knowledge and practice: Teacher learning in communities. In Iran-Nejad, A. & Pearson, C.D. (eds.), *Review of research in education*, (Vol. 24, pp. 249-305). Washington DC: American Educational Research Association.
- Doyle, W. (1990). Themes in teacher education research. In Houston, W.R. (ed.), *Handbook of research on teacher education* (pp. 3-24). New York: Macmillan.
- Dwyer, C. A., & Stufflebeam, D. (1996). Teacher Evaluation. In Berliner, D.C. & Calfee, R. C. (eds.), *Handbook of educational psychology* (pp. 765 - 786). New York: Macmillan.
- Freeman, D. & Richards, J. C. (eds.) (1996). *Teacher learning in language teaching*. Cambridge: CUP.
- Haertel, E. H. (1991). New forms of teacher assessment. In Grant, G. (ed.), *Review of research in education*, 17, 3-29. Washington: American Educational Research Association.
- Meijer, P. (1999). *Teachers' practical knowledge: Teaching reading comprehension in secondary education*. Academisch proefschrift. Leiden: Leiden University.
- Raamplan voor de Universitaire Lerarenopleiding. (1996). Utrecht: VSNU.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1- 22.
- Stuurgroep beroepskwaliteit leraarschap (1997). *Tekenen voor kwaliteit*. Groningen: Wolters-Noordhoff.
- VSLPC (1999). *Startbekwaamheden leraar secundair onderwijs*. Meppel: Ten Brink.
- Verloop, N. (1994). *Oriënterende studie naar de mogelijkheid tot ontwikkeling van een beoordelingsstelsel voor docenten*. Leiden: ICLON.
- Verloop, N., Beijaard, D., & Driel, J. van (1998). Beoordeling van leraren. In Verloop, N. (red.), *75 jaar onderwijs en opvoeding: 75 jaar Pedagogische Studiën* (pp. 171-187).
- Woods, D. (1996). *Teacher cognition in language teaching: Beliefs, decision-making and classroom practice*. Cambridge: CUP.