

Thema

**Passen en meten:
naar deugdelijk assessment
in organisaties en onderwijs**

GASTREDACTIE:

**ANS GROTENDORST
HARM TILLEMA**

REDACTIE:

**HANS DEKKER
SASKIA TJEPKEMA
ROB POELL
SIBRENNE WAGENAAR**

HRD Thema verschijnt 4 x per jaar
Jaargang 3, nr. 4 2002
ISSN: 1566-757X

Redactie

Hans Dekker, Rob Poell, Saskia Tjepkema,
Sibrenne Wagenaar

Productie

LINE UP tekstproducties bv, Groningen

Uitgever

Jacqueline Vreeburg

Abonnementen
Klantenservice HRM
Tel. 0570-673357
info@kluwer.nl

Kluwer
Postbus 23
7400 GA Deventer

Nederlands
uitgeversverbond
**Groep uitgevers voor
vak en wetenschap**

KLUWER 



Voorwoord



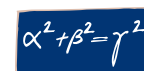
Tot slot



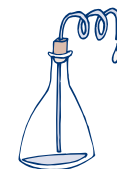
Onderzoek



Praktijk



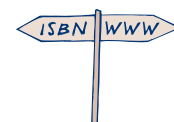
Theorie



Methode



Opinie



Literatuur- en
Webwijzer

INHOUD

VOORWOORD

DEUGDELIJK ASSESSMENT: PASSEN EN METEN

HANS DEKKER, ANS GROTENDORST EN HARM TILLEMA

5



WAT IS EEN ASSESSMENT CENTER?

DE ASSESSMENT-CENTERMETHODE, TOEPASSINGEN EN TECHNOLOGIEËN

WILLIAM C. BYHAM

7



In dit vertaalde artikel wordt ingegaan op de werkwijze van een assessment center, de validiteit en rechtvaardigheid ervan, het gebruik van de methode buiten de VS en op de nieuwe simulaties, testen en methoden die in de loop der tijd aan het assessment center zijn toegevoegd.

EEN TOETSSTEEN VOOR ASSESSMENTS

HARM TILLEMA

15



Deze bijdrage bespreekt de kwaliteitsbepaling van assessments, waarbij drie zaken centraal staan: de kwaliteitskeuzen die aan de orde komen bij de beslissing tot inzet van assessments in een organisatie, de kwaliteitsoverwegingen bij de feitelijke vormgeving van assessments en de kwaliteitsbeoordeling rond het afnemen en onderhouden van assessments.

EXAMENS: SLEUTEL, GEEN SLOT.

OP WEG NAAR NIEUWE EXAMENVORMEN IN HET (BEROEPS)ONDERWIJS

ANS GROTENDORST EN SASKIA SCHENNING

28



In de beeldvorming en berichtgeving over examens voeren klassieke vormen vaak de boventoon: op leerstof- en feitenkennis gerichte toetsing, die tijd- en locatiegebonden is en massaal. Volgens de auteurs verdient dit beeld zo langzamerhand wel de nodige nuanceering, aangezien in alle onderwijssectoren het denkproces over nieuwe examenvormen in volle gang is.

COMPETENT BLAUW OP STRAAT.

PROEVEN VAN BEKWAAMHEID IN HET VERNIEUWDE POLITIEONDERWIJS

BART BOLHUIS EN RICHARD VAN ROYEN

36



Het LSOP Politie Onderwijs- en Kenniscentrum is in 2002 gestart met vernieuwd politieonderwijs, waarbij de verwerving van competenties centraal staat. Hierdoor is examineren gericht op het meten van kennis en inzicht grotendeels vervangen door het beoordelen van bekwaamheden door middel van verschillende examenopdrachten. De auteurs beschrijven hoe zij zorgdragen voor 'deugdelijk assessment' door middel van proeven van bekwaamheid.

KANTTEKENINGEN BIJ EEN HYPE.

HOGESCHOLEN HEBBEN WEINIG OOG VOOR DE CONSEQUENTIES VAN INTAKE-ASSESSMENTS

ANJA GELDERMANS

48



Dit artikel doet verslag van een gesprek tussen opleiders over het gebruik van intake-assessments door hogescholen.

ADVISERING RONDOM SUCCES: VERANTWOORDELIJKHEDEN, VALKUILEN EN MOGELIJKHEDEN MET BETREKKING TOT DE TOEPASSING VAN ASSESSMENTS

ANTHON VAN DER HORST

Deze bijdrage gaat onder meer in op de verantwoordelijkheden die het beoordelen van het functioneren van mensen in arbeidsrelaties met zich meebrengt voor de opdrachtgever, de deelnemer en de beoordelaar.

52



HET DESIGNPROCES VAN ASSESSMENT CENTERS: SCHIPPEREN TUSSEN PRAKTIJK EN KWALITEIT

FILIP LIEVENS EN NANJA KOLK

In een assessment center worden belangrijke beslissingen genomen op basis van iemands prestatie. De kwaliteit van deze beslissingen was en is onderwerp van talloze studies. De aanbevelingen die hieruit volgen, moeten in de praktijk echter ook implementeerbaar zijn. In dit artikel worden deze twee perspectieven belicht, met als doel de kloof tussen wetenschap en praktijk te overbruggen.

63



BEOORDELEN IN COMPETENTIEGERICHTE LEEROMGEVINGEN

MARIANNE ELSHOUT-MOHR, RON OOSTDAM EN MARIANNE OVERMAAT

Een onderwerp dat op dit moment (weer) sterk in de aandacht staat, is de aansluiting tussen onderwijs en arbeid. Deze bijdrage formuleert een antwoord op de vraag hoe opleidingen studenten kunnen afleveren die over de competenties beschikken die de arbeidsmarkt van beginnende 'professionals' vraagt.

71



360°-FEEDBACK ALS HULPMIDDEL VOOR HET EVALUEREN VAN OPLEIDING EN ONTWIKKELING BINNEN BEDRIJFSORGANISATIES

FROUKJE JELLEMA

Organisaties hebben in toenemende mate behoefte aan inzicht in de effecten van opleiding en ontwikkeling van werknemers. Er is echter een tekort aan praktische en bruikbare methoden die zich hierop richten. Dit artikel gaat in op het gebruik van 360°-feedback als methode voor het evalueren van veranderd werkgedrag als gevolg van opleiding en ontwikkeling.

82



TOT SLOT TERUGKIJKEN OF VOORUITBLIKKEN?

ANS GROTENDORST, HARM TILLEMA EN HANS DEKKER

90



LITERATUUR- EN WEBWIJZER

ANS GROTENDORST, HARM TILLEMA EN HANS DEKKER

96



OVER DE AUTEURS

100



VOORWOORD

DEUGDELIJK ASSESSMENT: PASSEN EN METEN

HANS DEKKER, ANS GROTENDORST EN HARM TILLEMA

ASSESSMENT

Dat is waar we het in dit *HRD Thema* over hebben: assessment. De term ‘assessment’ komen we steeds meer tegen, in vakliteratuur maar ook zeker in de praktijk. Had assessment tien, twintig jaar geleden een zuiver beoordelingskarakter (denk aan werving en selectie: vaststellen van geschiktheid voor een beoogde functie), nu zien we dat het begrip een bredere betekenis heeft gekregen, en een bredere toepassing. In competentiegerichte werk- en leeromgevingen heeft assessment tegenwoordig de belangrijke functie om gedrag en prestaties van mensen zichtbaar te maken. Enerzijds met als doel het functioneren van medewerkers te kunnen beoordelen en een passende (aan)sturing van personeel te bepalen. Anderzijds om feedback te geven, waardoor medewerkers zich verder kunnen ontwikkelen. De bijdragen met de icoon ‘theorie’ gaan verder in op de conceptuele achtergronden van dit fenomeen.

DE ASSESSMENTPRAKTIJK

De interesse om via assessment competenties vast te stellen is groot, zowel in organisaties als in (beroepsgerichte) onderwijsinstellingen. Beide contexten vindt u dan ook terug in deze uitgave. De afgelopen tijd is er veel aandacht geweest voor de ontwikkeling van nieuwe assessmentinstrumenten en de toepassing

ervan. Te denken valt aan assessment centers, portfolio's en 360°-feedback. In de bijdragen die zijn voorzien van de icoon ‘praktijk’ en ‘methode’ leest u hier meer over.

KRITISCHE VRAGEN OVER DEUGDELIJKHEID: PASSEN EN METEN

Veel ervaringen met assessment leiden tot enthousiasme. Aan de andere kant (of daardoor?) is er tegelijkertijd een andere stroom waar te nemen. Een stroom die steeds nadrukkelijker kritische vragen stelt over de kwaliteit van de gehanteerde methoden; die zich richt op het vraagstuk van deugdelijkheid. Dit ‘accountabilityperspectief’ wint aan terrein omdat steeds duidelijker wordt dat assessmentpraktijken niet alleen zwaarwegende gevolgen hebben voor degene die onderworpen wordt aan een assessment, maar van organisaties ook een aanzienlijke investering in tijd, infrastructuur en faciliteiten vragen. Dan ligt het voor de hand dat het op z'n minst aannemelijk is dat de uitkomsten van deze praktijken valide en betrouwbaar zijn. Als gevolg hiervan komt er ook meer onderzoek van de grond dat zich richt op de kwaliteit of deugdelijkheid van assessment. Zoals we ook in deze aflevering van *HRD Thema* zien, levert dit kwaliteitsvraagstuk een continu passen en meten op tussen enerzijds willen voldoen aan wetenschappelijk verantwoorde eisen,

en anderzijds kiezen voor datgene wat mogelijk is binnen de praktijk. In de bijdragen die zijn aangeduid als ‘theorie’ en ‘opinie’ en ‘onderzoek’ zijn de kritische noten over deugdelijkheid te vinden.

eerst uw eigen conclusies trekken, om achteraf vast te stellen of te kijken of u tot dezelfde of andere inzichten komt dan de gastredactie.



DILLEMA'S: PASSEN EN METEN

Naast oplossingen brengt assessment ook nieuwe dilemma's met zich mee, die eveneens leiden tot passen en meten. Bijvoorbeeld: in steeds meer situaties hangt het vaststellen van bekwaamheid samen met het verwerven van bepaalde bevoegdheden, zelfs met certificering. Hoe kun je in zo'n context het ontwikkelingsgerichte perspectief hanteren? Wanneer houdt het leergerichte perspectief op en begint het beoordelingsperspectief? Een ander groot dilemma is het feit dat assessment inhoudt dat iemands competenties aan een standaard afgemeten worden. Maar wie is daarin normstellend, zeker gezien vanuit het ontwikkelingsperspectief? Dit type vraagstukken komt met name aan bod in de bijdragen met de icoon 'theorie' en 'opinie'.

CENTRALE VRAGEN

Hierboven gaven we een eerste beeld van de focus van dit *HRD Thema* op assessment. Samenvattend schetsen we hieronder de centrale vragen die we via dit themanummer willen beantwoorden:

- Wat kunnen we verstaan onder assessment, wat is het verschil met het gebruikelijke beoordelen? Waar komt de huidige belangstelling voor dit fenomeen vandaan?
- Wat zijn de overeenkomsten en benaderingen en toepassingen die men kiest in bedrijfsleven en onderwijs? En wat zijn de eventuele verschillen? Hoe laten die zich verklaren?
- Zijn we met assessment op het goede, deugdelijke spoor? Komen de essentiële zaken voldoende aan bod of vergeten we belangrijke punten? Hoe zal assessment zich in de toekomst ontwikkelen?
- Assessment kent een ontwikkelingsperspectief, maar het gaat toch ook vaak om beoordelen. Hoe verhouden deze perspectieven zich tot elkaar?
- Bij assessment gaat het om het 'spiegelen' van iemands competenties aan een standaard of norm. Maar: wie bepaalt die standaard?

Om antwoorden op deze kernvragen te krijgen kunt u natuurlijk rechtstreeks doorsteken naar 'Tot slot'. Op die plaats ontvouwt de gastredactie namelijk haar versie van de antwoorden op deze vragen. Spannender is het als u eerst de afzonderlijke bijdragen leest (zie de inhoudsopgave voor een nadere typering hiervan). Elke auteur gaat in zijn/haar bijdrage op meerdere van deze kernvragen in. Op grond daarvan kunt u ook

$$\alpha^2 + \beta^2 = \gamma^2$$

WAT IS EEN ASSESSMENT CENTER?

DE ASSESSMENT-CENTERMETHODE, TOEPASSINGEN EN TECHNOLOGIEËN¹

WILLIAM C. BYHAM, PH.D.

DEVELOPMENT DIMENSIONS INTERNATIONAL

WERKWIJZE VAN EEN ASSESSMENT CENTER

De assessment-centermethode omvat meerdere beoordelingstechnieken, waaronder verschillende soorten functiegerelateerde simulaties, en soms gesprekken en psychologische testen. Praktijksimulaties die vaak in assessment centers worden gebruikt, zijn:

- postbakoefeningen;
- groepsdiscussies;
- simulatie van gesprekken met ‘ondergeschikten’ of ‘klanten’;
- fact-findingoefeningen;
- analyse/besluitvormingsproblemen;
- oefeningen in mondelinge presentatie;
- oefeningen in schriftelijke communicatie.

Het doel van simulaties is om gedrag op te roepen dat essentieel is voor de belangrijkste aspecten van de functie waarvoor de kandidaat wordt beoordeeld. Deze aspecten, ook wel dimensies of competenties genoemd, worden voorafgaand aan het assessment center gedefinieerd door de beoogde functie te analyseren. Bij deze functieanalyse worden het gedrag, de motivatie en de kennis vastgesteld die essentieel zijn om de desbetreffende functie goed uit kunnen te oefenen. Tijdens het assessment blijkt uit de praktijksimulaties hoe de kandidaat zich in de doelsituatie zal gedragen en wat zijn of haar kennis en vaardigheden zijn.

Een traditioneel assessment center duurt één tot drie dagen en is doorgaans voor zes deelnemers. Tijdens de simulaties worden de deelnemers geobserveerd door assessoren (meestal drie lijnmanagers) die getraind zijn in het observeren en beoordelen van gedrag en kennisniveau. De assessoren observeren bij elke simulatie andere deelnemers en maken aantekeningen op speciale observatieformulieren. Nadat de deelnemers alle simulaties hebben uitgevoerd, bespreken de assessoren gedurende één of meer dagen hun bevindingen met elkaar en komen ze tot een gezamenlijke beoordeling. Eventuele test- en gespreksresultaten worden bij de besluitvorming meegenomen. Het eindoordeel van de assessoren is een schriftelijk rapport over de sterke en zwakke punten van de deelnemer. Afhankelijk van het doel van het assessment center kan er tevens in het rapport staan vermeld hoe de deelnemer naar verwachting in de beoogde functie zal functioneren.

Het belangrijkste aspect van de assessment-centermethode is waarschijnlijk dat deze niet aan het huidige functioneren is gerelateerd, maar juist aan het toekomstig functioneren. Door te observeren hoe een deelnemer de problemen en uitdagingen van de doel functie of het doelniveau aanpakt (zoals dat in de oefeningen wordt gesimuleerd), krijgen assessoren een duidelijk beeld van hoe die persoon in de doelpositie zou functioneren. Dat is met name nuttig bij het



William C. Byham

1 beoordelen van personen die in hun huidige baan niet de gelegenheid hebben te laten zien dat zij over de benodigde capaciteiten beschikken voor de beoogde positie of functie. Dit is vaak het geval bij werknemers die een leidinggevende functie ambiëren, maar in hun huidige functie niet kunnen laten zien dat zij over leidinggevende capaciteiten beschikken.

Een organisatie die gebruikmaakt van een assessment center heeft niet alleen het voordeel van een nauwkeurigere diagnose en selectie, maar ook enkele indirecte voordelen. Kandidaten accepteren de rechtvaardigheid en nauwkeurigheid van promotiebeslissingen eerder en hebben een beter inzicht in de functie-eisen. Wanneer managers tot assessor worden opgeleid, verbeteren ze bovendien hun vaardigheden in vele andere managementtaken, zoals het afhandelen van prestatiebeoordelingen en het voeren van coachings- en feedbackgesprekken.

VALIDITEIT EN RECHTVAARDIGHEID

De assessment-centermethode zoals we die nu kennen, vindt haar oorsprong in de AT&T Management Progress Study (Bray, Campbell en Grant, 1974)². Bij dit onderzoek, dat eind jaren vijftig van de vorige eeuw begon, werden personen beoordeeld die in een managementfunctie bij een van de vestigingen van Bell Telephone begonnen; vervolgens werden zij gedurende de rest van hun loopbaan gevolgd. Dit onderzoek was ongebruikelijk, in die zin dat het pure research was. Noch de personen die werden beoordeeld, noch hun chefs werden op de hoogte gesteld van hun prestaties

in het assessment center. Bovendien mocht deze informatie geen invloed hebben op de loopbaan van de deelnemers. De beoordeling van de deelnemers vond plaats vlak nadat ze als pas afgestudeerde in een managementfunctie waren begonnen of nadat ze van gewoon werknemer tot leidinggevende waren gepromoveerd.

De onderzoekers volgden niet alleen hoe de loopbaan van de deelnemers zich in de daaropvolgende jaren ontwikkelde, maar acht jaar na het eerste assessment werd ook nog een tweede assessment uitgevoerd (Howard en Bray, 1988). Het gebruikte criterium was doorstroming naar het vierde managementniveau in een hiërarchie van zeven niveaus. De voorspelling over acht jaar is steekhoudender – een verwachte bevinding aangezien de meeste deelnemers hun managementvaardigheden na acht jaar wel zouden zijn gaan consolideren. Desondanks waren de oorspronkelijke assessmentresultaten nog steeds van toepassing – zelfs na twintig jaar.

Thornton en Byham (1982) onderzochten 29 studies naar de validiteit van de assessment-centermethodologie. De auteurs vonden meer steun voor de assessment-centermethode dan voor andere selectiemethodologieën en betreurden het tegelijkertijd dat de meeste studies door maar een paar grote organisaties waren gedaan (AT&T, GE, IBM, SOHIO en Scars).

In 1985 verwerkten Thornton en diens medewerkers aan de Colorado State University 220 validiteitscoëfficiënten van vijftig studies met behulp van een statistische benadering, genaamd meta-analyse. Ze schatten de validiteit van de methode in op .37 (Gaugler, Rosenthal, Thornton en Bentson, 1985). Wayne Cascio van de University of Colorado werkte onafhankelijk van Thornton, maar kwam bij zijn onderzoek naar de validiteit van eerstelijns assessment centers in een vestiging van Bell System op dezelfde waarde uit (.37). Cascio was echter voornamelijk geïnteresseerd in het meten van de 'bottom-line impact' van promotiebeslissingen op basis van informatie uit assessment centers versus beslissingen die waren gebaseerd op criteria uit andere methoden (Cascio en Ramos, 1984).

Om de financiële impact van assessment centers te bepalen had Cascio meer nodig dan alleen validiteitsinformatie; hij moest kostengegevens hebben (volledig doorbelaste kosten van het assessmentproces) en gegevens met betrekking tot de prestaties, uitgedrukt in dollars. In een periode van vier jaar ontwikkelde hij een eenvoudige methodologie om de prestatieniveaus van managers in dollars uit te drukken. Cascio gebruikte de informatie die hem door meer dan 700 lijnmanagers werd aangereikt en combineerde gegevens over de validiteit en kosten van het assessment center met de in dollars uitgedrukte prestaties van eerstelijns managers. Op basis van deze gegevens maakte hij een

schatting van de nettowinst in dollars die de organisatie zou behalen door in het promotieproces gebruik te maken van assessment-centerinformatie. Over een periode van vier jaar werd de winst voor het bedrijf in termen van verbeterde prestaties van nieuwe managers geschat op \$ 13,4 miljoen of circa \$ 2.700 per jaar voor elk van de 1100 medewerkers die doorstroonden naar eerstelijns managementfuncties.

Ongunstige impact

Vergeleken met andere selectiemethodologieën wordt de assessment-centermethode over het algemeen als rechtvaardiger en objectiever beschouwd dan andere methodologieën wanneer het gaat om geslacht, ras en leeftijd.

Er zijn wel enkele differentiële prestaties gevonden, maar dat is doorgaans het gevolg van differentiële deelnemergroepen. De methode is uitzonderlijk rechtvaardig dankzij de nadruk op feitelijk gedrag in plaats van op psychologische constructen.

Consistent onderzoek toont aan dat assessment centers onbevooroordeeld zijn in hun voorspellingen van toekomstige prestaties. Bij deze studies werd gekeken naar leeftijd, ras en geslacht van een kandidaat en kon er worden geconcludeerd dat voorspellingen volgens de assessment-centermethodologie even betrouwbaar waren voor alle kandidaten. (Zie Thornton en Byham, 1982, voor een uitgebreide bespreking.)

Federale rechtbanken hebben geoordeeld dat assessment centers betrouwbaar en rechtvaardig zijn. Ze hebben zelfs meerdere malen opdracht gegeven tot het gebruik van assessment centers om selectieproblemen te verhelpen die voortkwamen uit het gebruik van schriftelijke testen en andere selectie-instrumenten.

GEBRUIK VAN DE ASSESSMENT-CENTERMETHODE BUITEN DE VERENIGDE STATEN

De belangrijkste reden waarom de assessment-centermethode in zo veel andere landen wordt gebruikt, is dat het een eenvoudig aan te passen beoordelingssysteem is en geen beoordelingsinstrument. Gebruikers hoeven geen dimensies of prestatienormen over te nemen die van belang zijn in de Verenigde Staten, maar in hun land wellicht totaal onbelangrijk zijn. Ze nemen alleen een systematische procedure over waarmee kandidaten kunnen worden beoordeeld op basis van functiegerelateerde dimensies die specifiek zijn voor hun organisatie en omgeving. De dimensie 'interpersoonlijke gevoeligheid' wordt in Japan bijvoorbeeld op veel meer manieren geuit dan in de Verenigde Staten, maar de methode waarmee deze dimensie wordt beoordeeld, werkt in beide landen gelijk (en even goed).

Selectie en benoeming van kandidaten voor hogere managementniveaus

Begin jaren zeventig van de vorige eeuw gingen organisaties de assessment-centermethode gebruiken voor de selectie en benoeming van hogere managers. Assessment centers zijn wel gebruikt voor de beoordeling van directeurs van bedrijven, fabrieksmanagers, algemeen managers en een groot aantal hogere overheidsfunctionarissen. Deze beoordelingen werden meestal verricht door een groep externe, 'professionele' assessoren (consultants). Het is moeilijk om binnen het bedrijf zelf gekwalificeerde mensen van hoog niveau te vinden die de tijd hebben om kandidaten objectief te beoordelen.

Selectie en benoeming van geautoriseerd personeel

De voornaamste groei van assessment centers sinds 1985 is gestimuleerd door organisaties die kiezen voor empowerment van hun medewerkers. Deze organisaties geven hun medewerkers:

- verantwoordelijkheid voor bepaalde gebieden of resultaten;
- beschikking over middelen, systemen, methoden en apparatuur;
- controle over werkomstandigheden en werktijden;
- beslissingsbevoegdheid (tot op zekere hoogte);
- beoordeling op basis van behaalde resultaten.

Bij de meeste organisaties maken de werknemers bovendien deel uit van zelfsturende teams. De teams bestaan uit teamleden en een teamleider (de teamleider is een meewerkend teamlid dat geen deel uitmaakt van het management). Het team is verantwoordelijk voor:

- verbetering van kwaliteit en productiviteit, job rotation;
- planning;
- wie werkt aan wat;
- kwaliteitscontrole;
- aanpassing, onderhoud en reparatie van materiaal;
- 'housekeeping', vakantieplanning, arbeidsverzuim, te laat komen en functioneringskwesties;
- kiezen van een teamleider;
- vele andere gebieden.

De introductie van zelfsturende teams verandert de rol van chefs en managers drastisch. Chefs (vaak ook groepsleiders genoemd) hebben een grote 'span of control', met soms wel honderd ondergeschikten. Omdat teams en teamleiders een groot deel van de gewone controlefuncties overnemen, krijgt de chef steeds meer een managementfunctie, met taken als budgettering en planning. Dat heeft op zijn beurt weer gevolgen voor het middenkader. De wijzigingen in

functies op meerdere niveaus hebben organisaties ertoe gedwongen nieuwe methoden te gebruiken voor het nemen van beslissingen rond selectie, promotie en benoeming. Omdat assessment centers zo goed werken op toezichthoudend en leidinggevend niveau, sprak het voor zich het assessment center als methodologie te gaan gebruiken.

Honderden productiebedrijven hebben gebruikgemaakt van assessment centers om werknemers, teamleiders en groepsleiders te selecteren. Om dat te realiseren werden er veel nieuwe processen ontwikkeld, met name voor het opstarten van 'greenfield' fabrieken waar grote aantallen sollicitanten moeten worden verwerkt. Toyota beoordeelde 22.000 sollicitanten om uiteindelijk de benodigde 3000 medewerkers voor de fabriek in Kentucky aan te stellen.

Op werknemerniveau moeten kandidaten in groeps oefeningen problemen oplossen, zijn er simulaties van het productieproces en één-op-één oefeningen in interactie. Oefeningen voor chefs bieden de gelegenheid vaardigheden te laten zien op het gebied van begeleiding, leidinggeven en besluitvorming.

Vaststellen van de trainings- en ontwikkelingsbehoeften

Snelle en eenvoudige trainingsmethoden veranderen het vaardigheidsniveau van mensen niet. Voor het aanleren van vaardigheden is intensieve, tijdrovende klaslokaaltraining nodig, die bovendien wordt gecombineerd met mogelijkheden tot on-the-job oefening en feedback zodat nieuw gedrag er als het ware inslijt. Omdat het ontwikkelen van vaardigheden veel tijd en moeite vergt, kan niet iedereen in elke vaardigheid

Vrijwel alle organisaties die voor selectie of promotie gebruikmaken van assessment centers, gebruiken de daarmee verkregen informatie ook om de trainingsbehoefte vast te stellen. Er is echter een duidelijke verschuiving merkbaar: een groot aantal bedrijven gebruikt assessment centers tegenwoordig alleen maar om vast te stellen welke training mensen nodig hebben. Dergelijke informatie kan medewerkers en organisaties veel tijd en moeite besparen doordat de personen in kwestie op het juiste moment het juiste trainingsprogramma wordt aangeboden.

Een diagnostisch assessment voor managers vindt meestal plaats binnen de betreffende organisatie met behulp van consultants. Assessment op directieniveau vindt plaats in uitgebreide, speciaal daarvoor opgezette assessmentfaciliteiten, die worden gerund door consultants.

Vaststellen van managementvaardigheden en vooronderstellingen als onderdeel van een veranderingsstrategie voor de bedrijfscultuur

Assessment van medewerkers in een fabriek of op een afdeling kan een integraal onderdeel van de veranderingsstrategie voor de bedrijfscultuur worden gemaakt. Nadat een organisatie voor een bepaalde cultuur heeft gekozen, is de volgende logische stap te bepalen welk gedrag nodig is om die cultuur te implementeren en te beoordelen wat de vaardigheidsniveaus van de functionarissen op dat gebied zijn. Een essentieel onderdeel van een participatieve cultuur is bijvoorbeeld de vaardigheid een vergadering zo te leiden dat alle deelnemers hun zegje kunnen doen en het gevoel hebben dat ze bij de besluitvorming worden betrokken. Of iemand deze vaardigheid heeft, kan worden vastgesteld in een assessment center.

Dergelijke diagnostische informatie is uiterst nuttig bij het ontwikkelen van een strategie om de bedrijfscultuur te veranderen. Mensen die niet in staat zijn participierend leiding te geven, kunnen geen participatieve strategie implementeren, zelfs als ze dat zouden willen – ze moeten eerst hun basisvaardigheidsniveau verbeteren. Daarnaast is uit onderzoek gebleken dat de gemakkelijkste manier om iemands attitudes of basisveronderstellingen over mensen te veranderen is om eerst het gedrag van diegene te veranderen. Deze insteek is duidelijk anders dan de eerdere strategie, waarbij organisaties probeerden de attitudes van mensen te veranderen in de hoop dat dan ook hun gedrag zou veranderen. Met deze nieuwe strategie kunnen de mensen worden geïdentificeerd wier attitudes of vooronderstellingen over mensen uit de toon vallen bij de gewenste cultuur. Hun gedrag wordt gewijzigd door

De assessment-centermethode biedt een effectieve manier om de trainings- en ontwikkelingsbehoeften te bepalen

worden getraind. De assessment-centermethode biedt een effectieve manier om de trainings- en ontwikkelingsbehoeften te bepalen, waarna werknemers in het voor hen meest geschikte programma kunnen worden geplaatst.

De assessment-centermethode is een uitstekend diagnose-instrument, omdat de vaardigheden van een persoon in specifieke gebieden (dimensies) worden onderverdeeld en vervolgens naar specifieke voorbeelden van goed en slecht gedrag binnen elke dimensie wordt gezocht. Hierdoor kan de persoon in kwestie en zijn/haar chef nauwkeuriger bepalen welke training en ontwikkeling nodig zijn.

middel van een effectief trainings- en ontwikkelingsprogramma. Hierin komen hun attitudes en veronderstellingen aan de orde door verbeterd bedrag te belonen met een versteviging van hun positie. Na verloop van tijd realiseert het management de gewenste cultuuromslag zo in de hele organisatie.

De effectiviteit van trainingsprogramma's evalueren

De American Society for Training and Development (ASTD) schat dat Amerikaanse bedrijven jaarlijks meer dan zestig miljard dollar uitgeven aan training. Het snelst groeiende deel van dit bedrag wordt besteed aan verkoop-, middenkader- en managementtraining, hoewel de meeste bedrijven de effectiviteit van hun trainingsprogramma's nog niet goed hebben geëvalueerd. De assessment-centermethodologie is een uitstekende methode om de validiteit en effectiviteit van trainingsprogramma's te evalueren. Over het algemeen worden er drie onderzoeksmethoden gebruikt. Bij de eerste methode wordt een groep mensen getraind, terwijl een vergelijkbare groep geen training krijgt. Beide groepen doorlopen vervolgens een assessment center. Bij de tweede en derde methode wordt een groep mensen beoordeeld, getraind en vervolgens nogmaals beoordeeld.

Organisaties als SOHIO, Lukens Steel, AT&T, de New York Metropolitan Transit Authority en enkele tientallen middelbare en hogere opleidingen bedrijfskunde hebben de assessment-centertechnologie gebruikt om trainingsprogramma's te evalueren (Byham, 1982). De komst van videotecnologie, waarmee mensen relatief goedkoop kunnen worden beoordeeld, heeft de toepassing van de assessment-centermethodologie op dit gebied enorm doen groeien.

NIEUWE SIMULATIES, TESTEN EN METHODEN

Simulaties zoals postbakoefeningen, groepsdiscussies, managementspelen en analyseoefeningen, zoals beschreven in Byhams artikel in *Harvard Business Review* in 1970, zijn nog steeds de basis van de assessment-centermethodologie. Er zijn echter wel nieuwe oefeningen bijgekomen, waarvan met name de gesprekssimulatie van belang is. In deze oefening krijgt de deelnemer achtergrondinformatie over de reden voor een gesprek met iemand (een ondergeschikte, collega of klant) en persoonlijke informatie over de betreffende persoon. Nadat de deelnemer de tijd heeft gekregen zich voor te bereiden, speelt hij samen met iemand die daarin getraind is het gesprek na in een rollenspel. De gesprekspartner speelt een van tevoren

vastgestelde rol en geeft standaardreacties voor alle onderwerpen die aan de orde komen. Een getrainde assessor observeert het gedrag van de deelnemer.

Hoewel groepsoefeningen zonder leider nog steeds veel worden gebruikt om leiderschap te beoordelen, zijn één-op-één gesprekssimulaties steeds populairder geworden. Deze verschuiving weerspiegelt het algemene gevoel dat vaardigheden in leidinggeven aan individuen niet per se in correlatie staan met vaardigheden in leidinggeven aan groepen. Een andere reden voor deze verschuiving is dat mensen die dezelfde groepsoefening doen, deze oefening elk totaal anders

Een toenemend aantal organisaties heeft voor een 'totaalsimulatie'-benadering van assessment gekozen

kunnen ervaren. Groepsinteractie is afhankelijk van de aard van de mensen in de groep. Soms is er veel concurrentie binnen een groep, en soms willen de groepsleden juist graag samenwerken. De ene keer willen meerdere mensen de leiding nemen, de andere keer is er slechts één persoon die de leiding neemt. Door dit gebrek aan consistentie hebben met name organisaties die zich richten op kwesties rond gelijke kansen op werk gekozen voor de meer gestandaardiseerde gesprekssimulaties of verschillende vormen van groepsoefeningen.

Een toenemend aantal organisaties heeft voor een 'totaalsimulatie'-benadering van assessment gekozen. In plaats van dat ze afzonderlijke en los van elkaar staande oefeningen gebruiken, hebben deze organisaties hun oefeningen geïntegreerd in een vast scenario. Personages die in de postbakoefening voorkomen, komen later terug bij simulaties en kandidaten spelen gedurende het gehele assessment dezelfde rol.

Het gebruik van videobanden om gedrag te stimuleren

Een ontwikkeling van de laatste vijftien jaar is het gebruik van videobanden om het gedrag van de deelnemers te stimuleren. Een deelnemer krijgt een video te zien van een situatie waarmee hij of zij op de werkplek te maken krijgt (bijvoorbeeld een interactie met een ondergeschikte). De band wordt zo nu en dan stopgezet, waarna de deelnemer uit vier mogelijkheden moet kiezen van wat hij of zij kan doen of zeggen. Vervolgens wordt een score berekend op basis van de antwoorden van de deelnemer in een aantal van deze situaties. Het scoresysteem is op basis van een validiteitsonderzoek ontwikkeld.

Bij andere toepassingen wordt de videosimulatie gebruikt om mondelinge reacties te krijgen. Deze toepassing is objectiever omdat de deelnemer een eigen antwoord kan geven en niet uit de meerkeuzeantwoorden hoeft te kiezen.

Psychologische vragenlijsten en projectietesten

De oorspronkelijke AT&T research assessment centers maakten gebruik van psychologische vragenlijsten (bijvoorbeeld de Edwards Personal Preference Schedule) en projectietesten (bijvoorbeeld de Thematic Apperception Test) ter aanvulling op observaties van het gedrag van deelnemers tijdens de assessment center simulaties. AT&T liet deze instrumenten vallen nadat ze hun research assessment centers hadden omgezet naar operationele assessment centers, die werden gerund door de Bell-bedrijven. AT&T liet de testen om twee redenen vallen: (1) de operationele assessment centers zetten managers in als assessor (in plaats van psychologen) en (2) men zette in deze periode (de jaren zestig) vraagtekens bij het gebruik van instrumenten als schriftelijke testen, in verband met de mogelijk nadelige impact op zogeheten 'protected groups'. De meeste organisaties die voor de assessment-centermethodologie kozen, deden dat in navolging van AT&T, zodat ze zich meer gingen richten op gedragsoefeningen dan op schriftelijke testen. De meeste organisaties lieten zelfs de intelligentietesten vallen in verband met de bekende negatieve invloed op de zwarte bevolking. De Bell-bedrijven hielden deze testen echter aan.

Tegenwoordig worden intelligentietesten (algemenevaardigheidentesten) weer gebruikt in combinatie met assessment centers. Onderzoeksgegevens tonen aan dat de combinatie van intelligentiegegevens en gedragsobservaties een duidelijk betere manier is om mensen te beoordelen dan het gebruik van deze zaken afzonderlijk (Thornton en Byham, 1982). Het probleem met het gebruik van schriftelijke intelligentietesten en andere psychologische instrumenten is dat deze zeer nauwkeurig moeten worden gevalideerd en dat assessoren speciaal moeten zijn getraind in het interpreteren van gegevens en het integreren van die gegevens met gedragsgegevens.

Projectietesten worden vaak gebruikt bij assessment centers op directieniveau. Vragenlijsten om functie- en kandidaatsmotivatieprofielen op elkaar af te stemmen worden op alle niveaus gebruikt.

Meerdere perspectieven (zelfrapportage, beoordelingsinstrumenten voor leidinggevend, ondergeschikten en collega's)

Een assessment center biedt inzicht in veel dimensies van een functie maar doorgaans niet in alle belangrij-

ke dimensies. Dimensies als werknormen en inzet worden in assessment centers niet goed beoordeeld. Om deze hiaten te vullen en om meer inzicht te krijgen in dimensies die in assessment centers worden beoordeeld, vullen veel organisaties hun assessment centers aan met zelfrapportage en beoordelingen door de leidinggevende en collega's of ondergeschikten van de desbetreffende persoon. Over het algemeen krijgt een deelnemer zes vragenlijsten met de doeldimensies en definities. De deelnemer vult een van deze lijsten in en geeft de andere vijf aan zijn of haar leidinggevende, collega's of ondergeschikten. Alle vragenlijsten worden rechtstreeks naar een centrale locatie gestuurd waar een computer de gegevens samenvat en een rapport genereert.

De combinatie van assessment center, zelfrapportage en beoordeling door leidinggevende/collega/ondergeschikte heeft een grote impact op de deelnemers. De feedbackadviseur en de deelnemer kunnen de resultaten van de verschillende bronnen (zelf, anderen en assessment center) vergelijken en tegen elkaar afzetten. Op basis van deze kennis kunnen ze nauwkeuriger bepalen welke dimensies verder moeten worden ontwikkeld.

Nieuwe methodologieën

Het grootste nadeel van het gebruik van traditionele assessment centers is de hoeveelheid tijd die managers erin moeten steken. Voor een gemiddeld assessment center geldt dat een manager zijn/haar werk twee tot drie dagen moet laten liggen om deelnemers bij simulaties te observeren, waarna nog één tot twee dagen worden besteed aan besprekingen met andere assessoren om tot een eindbeoordeling te komen. Hoewel managers het belang van selectie- en promotiebeslissingen inzien, besteden ze vaak met tegenzin zo veel tijd aan assessment. Een bijkomend probleem is de formaliteit van het traditionele assessment center, waardoor het center al gauw een 'gebeurtenis' wordt. Dat kan verwachtingen scheppen en de aandacht vestigen op degene die wordt beoordeeld en op degenen die niet voor een assessment zijn uitgenodigd. Bovendien dwingt het traditionele assessment center organisaties ertoe mensen het proces in groepen te laten doorlopen; de methode is zinloos als er slechts twee kandidaten voor een functie zijn.

Deze bezwaren hebben het gebruik van de assessment-centermethode binnen sommige organisaties beperkt tot slechts enkele selectie- en promotiebeslissingen. Daardoor zijn veel belangrijke en effectieve toepassingen, zoals het vaststellen van trainingsbehoeften, slechts in beperkte mate gebruikt. Organisaties zien wel in dat het steeds belangrijker wordt om eerst nauwkeurig de trainingsbehoeften van werknemers vast te stellen voordat ze naar trainingsprogramma's worden

gestuurd, maar de toepassing van assessment centers voor ontwikkeling wordt bemoeilijkt door een gebrek aan assessoren – terwijl de assessment-centermethodologie voor veel posities het beste diagnostische instrument is. Managers zien het belang van grondige en nauwkeurige diagnoses in, maar maken met tegenzin de tijd vrij die nodig is om de uitstekende diagnoses te produceren die de assessment-centermethodologie oplevert.

De methode informeler maken

Enkele organisaties in onder andere de Verenigde Staten hebben de eerdergenoemde implementatieproblemen opgelost door de assessment centers minder formeel en strak te maken, terwijl de basiscomponenten die voor de validiteit zorgen worden gehandhaafd. Organisaties doen dit door de assessment-centermethode te integreren in de dagelijkse activiteiten binnen het bedrijf, in plaats van dat de managers naar een speciale locatie of een ‘center’ hoeven te gaan.

Degene die moet worden beoordeeld krijgt een lijst met managers die in de vacature moeten voorzien. Hij of zij maakt over een periode van een aantal weken zelf een afspraak met elk van deze managers. De managers in kwestie kunnen de tijd voor de oefeningen zo in hun andere werkzaamheden inpassen.

Tijdens deze bijeenkomsten laten de managers de deelnemer dezelfde praktijksimulaties doen die ook bij formele assessment centers worden gebruikt. De ene manager vraagt de deelnemer bijvoorbeeld waarom hij of zij bepaalde beslissingen heeft genomen bij de postbakoefening; een andere manager laat de deelnemer resultaten presenteren uit een analyse- en planningsoefening; een derde manager kan de deelnemer beoordelen in een één-op-één gesprek met een andere manager die de rol van ondergeschikte speelt.

Op een bepaald moment komen de managers (assessoren) bijeen voor een assessorbespreking die net zo werkt als een bespreking in een traditioneel assessment center. De assessoren geven voorbeelden van het gedrag van de deelnemer ter illustratie van hun oordeel over de dimensies die ze moesten observeren. Nadat ze alle observaties hebben besproken, komen de assessoren voor elke dimensie tot een gezamenlijk oordeel over de sterke en zwakke punten van de deelnemer. Als het assessment center als basis moet dienen voor een beslissing over selectie of promotie, vormen de assessoren zich daar vervolgens een oordeel over. Als het assessment als doel heeft te bepalen welke training de deelnemer nodig heeft, stellen de assessoren tot slot een profiel op van de sterke punten van de deelnemer en de gebieden waarin training noodzakelijk is.

Alle belangrijke componenten van de assessment-centermethode zijn aanwezig; meerdere praktijksimula-

ties; gebruik van bij simulaties geobserveerd gedrag om toekomstig gedrag in de beoogde functie te voorspellen; organisatie van geobserveerd gedrag rond functiegerelateerde dimensies; en een systematische gegevensintegratiesessie met meerdere assessoren die de deelnemers onafhankelijk in de simulaties hebben geobserveerd. De procedure is alleen veel minder formeel en star. Hierdoor kan zelfs de kleinste organisatie de assessment-centermethode toepassen om selectie/promotiebeslissingen te nemen.

Video gebruiken om gedrag te registreren

Een andere technologie die steeds meer aan populariteit wint, is het gebruik van videoapparatuur om het gedrag van de deelnemer op te nemen. In plaats van dat assessoren de deelnemers in simulaties observeren, wordt het gedrag van de deelnemers op video opgenomen. De videoband en de schriftelijke bijdragen van de deelnemer kunnen dan vrijwel overal naartoe worden gezonden en assessoren kunnen de opgenomen en geschreven prestaties bekijken en beoordelen wanneer het hen uitkomt. Nadat elke assessor de hem of haar toegewezen simulatie heeft geobserveerd en beoordeeld, kan een standaard gegevensintegratiesessie worden gehouden. Maar de gegevens kunnen ook met behulp van een speciaal systeem door de computer worden geïntegreerd.

Geautomatiseerde gegevensintegratie

Softwareprogramma's kunnen het assessmentproces versnellen; daarom zijn veel assessoren en gegevensbeheerders groot voorstander van het gebruik ervan. In een gewoon systeem voeren de assessoren hun observaties direct in de computer in. De computer sorteert het gedrag op dimensie en presenteert de informatie zo dat de assessor de verschillende dimensies gemakkelijker kan beoordelen. In een speciaal systeem

Softwareprogramma's kunnen het assessmentproces versnellen

controleert de computer vervolgens de beoordeling. Als de beoordeling van de computer verschilt van die van de assessor, bekijkt een tweede assessor de gegevens waarna hij of zij deze met de eerste assessor bespreekt. Samen komen ze tot een eindoordeel voor de oefening.

Op de integratiebijeenkomst worden de gedragsobservaties van alle oefeningen door de computer geïntegreerd en op een overzichtelijke manier gepresenteerd, zodat de assessor de gegevens eenvoudig kan analyseren en er een besluit kan worden genomen. In

sommige organisaties wordt de integratiebijeenkomst vervangen door een speciaal systeem. Deze mathematische gegevensintegratie is mogelijk dankzij de grote betrouwbaarheid van de dimensiebeoordelingen van de assessoroefening, waarbij een betrouwbaarheid van .90 en hoger niet ongebruikelijk is.

De computer drukt een gedetailleerd eindrapport af waarin dimensiescores met gedragsvoorbeelden staan. Het computersysteem reduceert de tijdsinvestering van de assessor met meer dan vijftig procent. Bovendien hoeft er veel minder tijd te worden besteed aan training van assessoren en bestuurders.

Gebruik van externe assessoren

Ten gevolge van inkrimping, enzovoort, zijn er minder middenmanagers die als assessor kunnen fungeren. Dit heeft ertoe geleid dat meer dan de helft van de Amerikaanse organisaties die gebruikmaken van assessment centers zijn toevlucht heeft genomen tot getrainde externe assessoren om de selectie en het beheer van kandidaten te verzorgen.



Noten

1. © Development Dimensions International, Inc., MCMXCVI. Alle rechten voorbehouden.
2. Deze vertaling bevat verscheidene bronverwijzingen maar geen literatuurlijst, aangezien deze ook in de brontekst ontbreekt.

Over de auteur: William C. Byham, Ph.D., is voorzitter en algemeen directeur van Development Dimensions International, een wereldwijd opererend HR-adviesbureau gespecialiseerd in het ontwerpen van selectiesystemen en het ontwikkelen van managementtalent.

$$\alpha^2 + \beta^2 = \gamma^2$$



EEN TOETSSTEEN VOOR ASSESSMENTS

HARM TILLEMA

Toetsen en beoordelen van competenties – assessment – speelt een steeds belangrijker rol in werk, opleiding en leren. Niet alleen de bepaling van eerder verworven competenties, maar ook de vaststelling van geschiktheid voor toekomstige taken vraagt om assessment. De steeds gevarieerder wordende samenstelling van eerder verworven vaardigheden en kennis én de voortschrijdende heterogenisering van afgeronde leerroutes (reguliere vooropleiding duaal, verkort, deeltijd, in-company of via training) verlangen een onafhankelijke en verantwoorde beoordeling van iemands kwaliteiten. Het inrichten van beoordelingsprocedures is een belangrijke zorg geworden voor zowel opleidingen, certificeringsinstanties als selectiebureaus en bedrijven.

Om helderheid te bieden over wat onder de kwaliteitsbepaling van assessments begrepen kan worden, zijn drie zaken te onderscheiden, die in deze bijdrage centraal staan:

- de kwaliteitskeuzen die aan de orde komen bij de beslissing tot inzet van assessments in een organisatie;
- de kwaliteitsoverwegingen bij de feitelijke vormgeving van assessments;
- de kwaliteitsbeoordeling rond het afnemen en onderhouden van assessments.

Voordat we uitvoerig op elk van deze drie kwaliteitsaspecten ingaan, volgt eerst een korte introductie tot assessment.

ASSESSMENT: INTRODUCTIE

Hoe definieer je assessment en welke verschillende doelen of functies kan assessment hebben? En hoe onderscheidt assessment zich van traditionelere beoordelingsmethoden? Dat zijn de centrale vragen in deze introductie.

Assessment: definitie en doel

Assessment dient in de kern om vast te stellen wat de kwaliteit is van prestaties. Een goede Nederlandse term voor assessment is er niet, maar wellicht komt meten van bekwaamheid of competentiebeoordeling het dichtst in de buurt. In het streven naar meer ‘accountability’, ofwel het afleggen van verantwoording over het beoordelen van kwaliteiten van mensen, is assessment een belangrijke manier om vast te stellen, te laten blijken als het ware, welk gedrag en welke prestaties functioneel zijn en met welke competenties deze gepaard gaan.

Assessment wordt het meest gebruikt voor intakebeoordeling (Seegers, 1986), waarbij de beslissing aan de orde is of men wel of niet voldoet aan gestelde criteria. Verwijzing en het al of niet toelaten tot een opleiding of trainingstrajecten is vaak het oogmerk. Assessment kan ook heel goed dienen als diagnose bij de ontwikkeling van competenties, als voortraject of tij-

dens een opleidingsprogramma. We zien daarmee twee belangrijke functies voor assessment. De eerste is op selectie en toelating gericht gebruik, waarbij men bekijkt wat iemand heeft gepresteerd. De tweede betreft op ontwikkeling gericht gebruik, waarbij men kijkt waartoe iemand in staat is en wat iemand zou kunnen bereiken. De eerste functie is retrospectief bedoeld, terugblikkend van aard; de andere is prospectief, vooruitkijkend.

Het voordeel van een koppeling van beide assessmentfuncties is dat er een verband wordt gelegd tussen verlangd functioneren enerzijds en feitelijke prestaties anderzijds, en wel op concreet gedragsniveau. Op die manier is direct inzichtelijk te maken wat de sterke en zwakke punten zijn waarover iemand beschikt. Het is dan maar een kleine stap om te kiezen voor aanvullende scholing of ontwikkelingsactiviteiten op maat (Jansen, 1989). Juist deze koppeling maakt assessment tot een belangrijk instrument in het strategisch kiezen van leerwegen (Tillema, 1996).

Kortom, met assessment wil men een nauwe relatie leggen tussen competentiebeoordeling en leerwegprogrammering voor een persoon. Assessment is dus een bijzondere vorm van meten, met als doelen:

- accurate informatie te verschaffen over gedrag en de uitkomsten daarvan, liefst op een zo concreet mogelijke basis;
- een indicatie te geven voor het benodigde ontwikkelingsproces, wil er sprake zijn van het opheffen van gevonden discrepanties tussen feitelijk en verlangd functioneren.

Assessment versus andere beoordelingsvormen

Assessment onderscheidt zich van andere manieren van evaluerend meten (Shulman, Haertel en Bird, 1988), zoals toelatingsonderzoek, sollicitatiegesprekken, controle op verworven bekwaamheden, dossiers van eerder functioneren. Bij assessment ligt de nadruk op de authenticiteit en realiteitswaarde van toetsituaties om vast te stellen wat iemand kan (Herriot, 1989). De afname van een assessment verloopt meestal via speciaal ontwikkelde proeven of bewijzen van kunnen waaraan rechtstreeks en op een zo levensecht mogelijke manier het vereiste functioneren valt af te meten. Dit in tegenstelling tot paper-en-pencil testen, vragenlijsten of rapporten waarin het gedrag achteraf wordt geïnterpreteerd (Lee en Beard, 1994). De uitkomst van een assessment moet een indicatie geven van de kwaliteit van de prestaties van een persoon, het niveau van beheersing dat is bereikt. Het criterium voor kwaliteit is ontleend aan benoemde (beroeps)competenties (Berryman, 1993). We kunnen dit vertalen in de eis dat assessments authentiek en passend moeten zijn ten opzichte van de uiteindelijke beroepssituatie

omschreven in termen van competenties (Lee en Beard, 1994).

Assessmentinstrumenten zijn, gelet op deze eis, meestal direct gerelateerd aan die beroepssituatie, maar dit hoeft niet noodzakelijk zo te zijn. Ze kunnen ook zodanig geconstrueerd zijn dat ze belangrijke of saillante aspecten van de latere beroepssituatie bevatten. Gesimuleerd of niet, een assessmentinstrument moet achterhalen wat een persoon kan laten zien over verlangd functioneren. Daarbij is niet alleen bedoeld op het gedrag in de beroepssituatie, maar ook op de bij dit gedrag behorende kennis en attitudes (Wiggins, 1989).

Het gevaar bij de meer gebruikelijke instrumenten voor het vaststellen van gedrag, zoals paper-en-pencil testen, vragenlijsten, toetsen, interviewgesprekken en dergelijke, is dat zij vooral meten wat een persoon weet, maar niet of nauwelijks wat een persoon kan (Shulman et al., 1998). Een ander verschil is dat assessment steeds een directe registratie van gedrag veronderstelt, zoals het demonstreren, het tonen van bekwaamheid, het geven van een proef, het leveren van werkexemplaren of een prestatieproduct. Het gaat er dan ook om die methoden te vinden die voor dat doel geschikt zijn. Daarbij valt te denken aan toepassingsoefeningen, simulaties en games, observaties en interviews tijdens taakuitvoering, gedragsobservatie, teaminteractie/rollenspel, portfolio's, exhibities, work-samples, enzovoort (zie Tillema [1993] voor een overzicht).

Voor assessment geldt dus het volgende:

- Er wordt op een authentieke manier gemeten, dat wil zeggen: assessment dient gerelateerd te zijn aan de beroepssituatie of, nog anders gesteld, dient competentiegericht te zijn.
- Er worden directe, gedragsgerichte registratiemethoden gebruikt om de kwaliteit of het niveau van het gedrag vast te stellen.

KWALITEITSKEUZEN BIJ DE INZET VAN ASSESSMENTS

Voordat kan worden gekeken naar de kwaliteit van assessmentafnames, is het eerst van belang duidelijkheid te krijgen over een aantal zaken die de inzet ervan betreffen. Zij vormen als het ware de bouwstenen voor een goede kwaliteit. We onderscheiden ten minste tien keuzemomenten waarop een beslissing wordt verlangd.

1. Onderscheid in soorten assessmentdoelen

Assessment kan worden gebruikt voor verschillende doelen, namelijk selectie, verwijzing en ontwikkeling,

zo is eerder betoogd. Dit heeft consequenties voor het gebruik en de inzet ervan. Zo is de vraag of men de resultaten wil gebruiken voor feedback naar personen, voor het afgeven van geschiktheidsverklaringen of de voortgangsbepaling in trainingsprogramma's. Gaat het daarbij om 'self appraisal' van functioneren ofwel om externe beoordeling door assessoren? Belangrijk is ook om vast te stellen of mensen achteraf inzage krijgen in het assessment en vooraf mogen weten waaraan zij 'onderworpen' worden. Bekendmaking van het doel van een assessment maakt dat de uitkomsten ervan beter bespreekbaar zijn, zeker waar het de te nemen vervolgacties betreft.

2. Inzet van assessment

Om gericht assessments te kunnen ontwerpen is het nodig nauwgezette afspraken te maken over de inzet van de in een assessment verkregen gegevens in de organisatie of opleiding. Dit kan bijvoorbeeld door verlangde eindniveaus vast te leggen, screenings uit te voeren van mogelijke kandidaten voor het assessment, gebruik van gegevens en verwachtingen van het assessment te omschrijven en records aan te leggen van additionele gegevens over functioneren, om zodoende het assessment zelf meer te kunnen richten op specifiek omschreven doelen.

3. Gebleken of verlangde geschiktheid

Het verzamelen van 'informatierecords' over gebleken competenties van personen is moeilijk. Hier doet zich namelijk een dilemma voor, want bij assessment wil men screenen op geschiktheid voor nieuwe of toekomstige competenties, terwijl gebleken geschiktheid daarvoor niet per se de ideale graadmeter hoeft te zijn. Enerzijds wil men huidig gedrag zo goed mogelijk waarderen om feitelijke start- of beginniveaus te distilleren, terwijl anderzijds men eigenlijk kijkt naar de eisen die in de voorgenomen leerweg een rol spelen. Om dit dilemma op te lossen kan men gebruikmaken van zo veel mogelijk gevarieerde bronnen: rapporten over eerder functioneren, 'bewijzen van bekwaamheid' zoals producten van werk, commentaren van direct betrokkenen uit vooropleidingen, enzovoort.

4. Afspraken over de afname

Om assessments af te nemen is het nodig methodisch, systematisch en controleerbaar te werken, niet alleen om ethische redenen maar ook om de betrouwbaarheid van gegevensverzameling te bevorderen. Zo zijn er bijvoorbeeld afspraken nodig over wie inhoudelijk beoordelaar/assessor is, wie de gegevens verzamelt en verwerkt, wie rapporteert, wie deel uitmaakt van de assessmentprocedure en of er sprake is van getrainde

assessoren. Andere zaken waarover afspraken moeten worden gemaakt, zijn hoe de verantwoordelijkheid is verdeeld, of er sprake is van een beroepsprocedure, hoe wordt teruggedoorgerapporteerd en hoe feedback wordt gegeven.

5. Voldoende informatie verzamelen

Bij een assessment zal men uiteindelijk een keuze willen maken uit de meest geschikte instrumenten. Een overweging daarbij kan zijn om een grote spreiding te hebben van instrumenten voor verschillende situaties; we spreken dan van een horizontale prestatievergelijking. Men kan zich ook willen richten op verticale prestatievergelijking, waarbij diepgaander wordt gemeten binnen eenzelfde taakgebied. Beide vormen vereisen andere en vaak meerdere assessmentinstrumenten. De moeilijkheid is erop te letten dat de beoordelingen onderling vergelijkbaar zijn. De overweging blijft steeds of men aanvullende gegevens alsnog moet betrekken in de eindbeoordeling; te denken valt bijvoorbeeld aan zelfbeoordeling. De ervaring leert dat dergelijke aanvullende gegevens het beste pas in een eindbeoordeling een rol kunnen spelen om te vermijden dat overwegingen een rol gaan spelen die weinig te maken hebben met een valide assessment.

6. Variatie in instrumenten

Dankzij de aanwezigheid van een grote hoeveelheid beoordelingsinformatie is een weloverwogen evaluatie mogelijk. In de assessmentliteratuur zijn vele beschrijvingen gegeven van geschikte methoden van assessment, zoals de assessment-centermethode, met daarin de in-basket-techniek, het rollenspel of simulatieoefeningen. Verder valt te denken aan het portfolio als dossier van kunnen of gerichte taakopdrachten als de probleemanalyse en andere, minder vaak genoemde, maar soms evenzeer geschikte methoden, zoals het tweegesprek of de kritieke-incidentenmethode. Zie Tillemans (1993) voor een overzicht van basisvormen.

7. Protocollen voor (controleerbare) meting

Authenticiteit van assessment, dat wil zeggen de verbinding met de beroepssituatie, leidt tot validiteitsverhoging, maar de betrouwbaarheid, ofwel de accuratesse en precisie van de meting, kan daarbij in gevaar komen. Het is daarom zaak in elk geval de manier waarop de beoordeling totstandkomt zo nauwgezet en controleerbaar mogelijk te maken. Dit kan bijvoorbeeld door het hanteren van registratieprotocollen door assessoren en explicitering van beslisriteria. Dit veronderstelt dat er van tevoren (handleidingen voor) competentieanalyses en gedragsspecificaties zijn gemaakt.

8. Standaarden vaststellen

De eis van (controleerbare) meting vooronderstelt het opstellen van gedragsspecificaties. De meest gebruikelijke gang van zaken om gedragsspecificaties te bepalen is om verlangde competenties uit te werken in indicatorenlijsten of in categorieënsystemen waarin typerende gedragingen zijn geclassificeerd, dan wel geordend naar niveau. De scoringsregels moeten echter expliciet zijn, wat minder gemakkelijk is dan het lijkt.

Bij een assessment is er steeds sprake van meerdere assessoren; zij verrichten de scoring van gedrag. Deze 'menselijke factor' is er de oorzaak van dat er, soms aanzienlijke, inter-assessorverschillen kunnen ontstaan bij de interpretatie van gedrag. Expliciete scoringsregels kunnen helpen voorkomen dat er te grote verschillen ontstaan. Meer standaardisering kan echter ten koste gaan van de authenticiteit van de meting.

9. Koppeling van assessment met ontwikkeling

Assessment past heel goed in een traject van opleiding en ontwikkeling. Bij ontwikkelingsgericht assessment laat men de beoordeling van gedrag uitmonden in een persoonlijk ontwikkelings- en opleidingsplan, van waaruit gerichte onderwijsactiviteiten zijn vorm te geven. Het is daarbij goed te bedenken dat het assessment op zichzelf eigenlijk al een opleidings- en leer-moment is op basis waarvan ontwikkeld kan worden (!) – de uitkomst van een assessment verwijst immers naar benodigde vervolgopleidingsactiviteiten. Er wordt in dit verband vaak gesproken van 'development centers', waarin de opgespoorde zwakke punten in iemands functioneren tot onderwerp van voortgezette scholing voor die persoon worden gemaakt.

10. Feedback

Door middel van assessment verkrijgt men beslissingsinformatie. Het gaat om informatie met een hoge sturingswaarde en de vraag is dan ook hoe deze assessmentinformatie wordt gebruikt. De feedbackinformatie leent zich voor processturing maar ook voor een succesindicatie over vertoond gedrag; het kan motiverend ofwel beoordelend worden gebruikt, en de feedback naar personen kan verwijzend of correctief van aard zijn.

KWALITEITSOVERWEGINGEN BIJ DE VORMGEVING VAN ASSESSMENTS

Een systematische aanpak om assessments te maken veronderstelt een proces dat ingaat op ten minste de volgende ontwerpcategorieën: nauwkeurige analyse van de (latere) beroepssituatie in termen van competenties, het maken van een assessmentprocedure

bestaande uit verschillende oefeningen/taken, training van assessoren in het beoordelen en evalueren ervan en afname van het assessment.

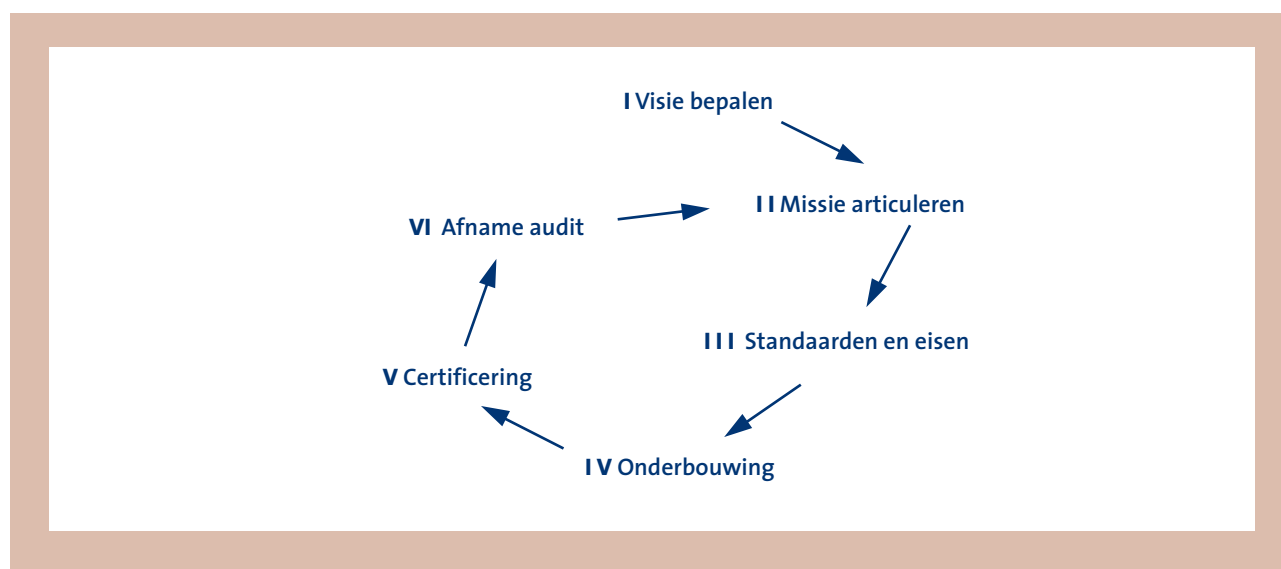
Een projectmatige opzet voor het ontwerpen van assessments ziet er mogelijk als volgt uit:

Voortraject

1. Er wordt een team samengesteld, bestaande uit vertegenwoordigers van het werkveld, samen met enkele opleidings-, leerwagspecialisten en, als eenmaal het domein/vakgebied is vastgesteld, terrein-deskundigen. Het projectteam omvat dus ter zake geïnformeerde personen. Dit projectteam wordt waar nodig aangevuld met assessmentdeskundigen/assessoren.
2. Het projectteam werkt zich inhoudelijk in op de randvoorwaarden en keuzen voor de kwaliteit van een assessment (zie de tien punten die hiervoor genoemd zijn). Het is van belang dat ieder projectlid op de hoogte is van deze keuzeoverwegingen, de beoogde output en daarbij tevens een 'sense of contextuality'/'gebruik' ontwikkelt.
3. De competentieanalyse verloopt in het team via a) een individuele bevrachingsronde en b) een daarop aansluitende groepsgewijze confrontatieronde, waarin kenmerkende situaties uit de beroepsuitoefening worden geïnventariseerd die passen bij de leerweg of het competentiedomein in kwestie. Ter aanvulling kan worden besloten enkele verkennende studies uit te voeren naar beschikbare kennis van personen die de beroepssituaties van 'binnenuit' kennen, om zo niet alleen gedragsaspecten maar ook andere kennis- en houdingsaspecten in het assessment te betrekken.

Uitvoeringstraject

4. De ontwikkeling van een assessmentprocedure verloopt langs twee paden. Enerzijds worden basisvormen van instrumenten (die de assessmentdeskundigen ter beschikking stellen) in het projectteam gescreend op hun waarde om in het assessment te gebruiken. Dit gebeurt door vast te leggen op welke aspecten men vooral wil beoordelen (levensechtheid, complexiteit, sociale aspecten van functioneren; zie Tillema [1996] voor een keuzematrix). Anderzijds wordt de keuze voor bepaalde instrumenten in zogenoemde panelronden aan elkaar voorgelegd en besproken aan de hand van handleidingen. Dit is een open proces met als doel het genereren van concreet uitgewerkte assessments, die vervolgens worden aangepast en verbeterd en uiteindelijk in afnameprocedures vastgelegd. Er zal in deze stap vaak sprake zijn van een behoefte aan afstemming over beoordelingscriteria, omdat de ontwikkeling van assessments discussie oproept



Figuur 1. Stappen in een auditproces

over al of niet stringente specificatie van criteria, meer dan vooraf vaak mogelijk is.

5. Nadat de assessmentinstrumenten zijn bepaald, volgt de codificering ervan, dat wil zeggen: het definitief vastleggen van informatie in protocollen en handleidingen omtrent de algemene beschrijving van het assessment, het doel en de competenties die worden gemeten, het verloop van het assessment, de taakstelling van de assessor, scoringsprocedures, eisen die worden gesteld aan assessoren en hun voortraining, benodigde materialen en ‘resources’ en punten van aandacht bij de afname.

Natraject

6. Bijstelling en revisie: via een duidelijk afsprakenstelsel worden bij (specifieke groepen) personen de assessments afgenomen. In het projectteam is reeds nauwkeurig nagegaan welke (groepen) personen daarvoor in aanmerking komen. Er kan sprake zijn van verschillende revisieronden voor het vaststellen van geschiktheid en kalibratie van een assessment.

KWALITEITSBEOORDELING BIJ DE AFNAME VAN ASSESSMENTS

Het uitvoeren van assessments vraagt om interne controle op kwaliteit, niet alleen om kwaliteit te waarborgen en de assessments te voorzien van een legitimering, maar ook om op een onafhankelijke manier te controleren of de afgegeven beoordelingen (bijvoorbeeld certificaten) de gewenste opbrengsten hebben. Een dergelijke controle valt aan te duiden met de term ‘audit’. Een audit kan verzekeren dat assessments voldoen aan gestelde eisen en biedt een garantie dat intenties ook daadwerkelijk worden waargemaakt. Een

audit van assessments hoort in feite onlosmakelijk bij een kwaliteitsinfrastructuur rond assessment.

Vanwege het grote belang van de controle op kwaliteit bij de afname van assessments gaan we uitvoeriger op een aantal zaken in.

Audit als een proces

Te gauw is men geneigd een audit af te doen als enkel een instrument; het is echter meer dan dat. De daadwerkelijke afname van een audit kan men wellicht beschouwen vanuit een instrumentele optiek, maar in feite gaat aan de afname van de audit een proces van kwaliteitsdenken en -afweging vooraf. In het schema in figuur 1 is nader uitgewerkt welke fasen in het kwaliteitsborgingsproces van de audit aan bod komen. Het schema geeft aan dat een audit nauw samenhangt met de verantwoording van het hele assessmenttraject en er ook daadwerkelijk mee verweven is.

De volgende vragen mogen verduidelijken dat men al in een vroeg stadium van ontwikkeling en ontwerp van assessments bezig moet zijn met het verzekeren van kwaliteit via gerichte auditvragen.

1. Een visie op assessment bepalen

Voordat men met assessments kan gaan werken en deze als onderdeel van het takenpakket gaat opnemen (bijvoorbeeld door middel van een eigen assessmentbureau), is nadrukkelijk eerst de eigen beleidskeuze aan de orde. De keuze voor assessment valt nader uit te werken in leidende thema's, de ‘shared vision’, rondom wat men met assessment in de organisatie of opleiding wil. De centrale vraag is waarvoor assessment moet dienen. Dit betreft uitwerking van een visie op

vragen als: wat betekent het als iemand bij ons een assessment doorloopt, welke bijdrage leveren wij met het hanteren van assessments, wat heeft men eraan als men bij ons een assessment heeft doorlopen?

Concreet zal een aantal eigen keuzen moeten worden gemaakt (en moeten die keuzen niet klakkeloos van anderen worden overgenomen):

- Wat willen we met assessment?
- Welke typen en vormen willen we aanbieden (ontwikkelingsgericht of geschiktheidsbeoordeling, breed of alleen voor intake)?
- Wat voor beeld houdt iemand over als het assessment doorlopen is?
- Hebben we een idee over leren en ontwikkelen door middel van het assessment, is er advisering mogelijk rond het vervolgtraject?
- Hoe is een en ander vanuit de literatuur of eigen kwaliteitsonderzoek onderbouwd? En hoe grondig is de aanpak gevalideerd of moet de aanpak gevalideerd zijn?
- Hoe opereren we in de markt, welke groepen willen we bedienen en welke samenwerking is daarbij nodig?

2. De missie articuleren

Een visie geeft de *raison d'être*, de reden om ergens mee door te gaan. Maar vervolgens is het nodig concreet aan te geven wat men met assessments wil, ingebed in de aanpak en werkwijze van de instelling als geheel. Men zou hier kunnen spreken van het businessplan rond assessment, met als centrale vraag daarbij: wat willen we in de instelling met assessment bereiken en wat moet het dan opleveren?

Auditvragen die rond deze articulering van functie en rendement zijn te beantwoorden, zijn:

- Welke specifieke doelen dient het assessment; wat doen we wel en wat niet met assessments?
- Wat is precies onze dienstverlening met het assessment en aan wie is die gericht?
- Welke concrete varianten en versies werken we in het assessment uit met welke diverse typen en vormen?
- Welk marktsegment bedienen we daarbij?
- Welke focus of nadruk is er aangebracht in het assessment? Willen we bepaalde aspecten meer benadrukken in het assessment of hebben we een generiek assessment voor ogen?
- Met welke infrastructuur en middelen inzet onderhouden en regelen we de assessments?

Met andere woorden: er dient een interne organisatie opgebouwd te worden waarbinnen activiteiten rond assessment kunnen worden ingebed.

3. De standaarden en eisen rond de opbouw van het instrumentarium omschrijven

Gelet op wat men met assessment beoogt (de vorige stap) is het vervolgens nodig de eigen kwaliteitsstandaarden aan te geven, de standaarden die men heeft voor het hanteren van assessments. De centrale vraag is hier: wat is onze toetssteen om kwaliteit van het assessment te bepalen en te waarborgen? Een belangrijk element, aangegeven in deze stap, is de kwaliteitsbepaling met betrekking tot de gehanteerde kwalificatieniveaus: tot welk niveau en tot welke graad van detaillering wil men kwaliteit zichtbaar maken, hoe goed wil men zijn (waarbij 'goed' een bepaalde prijs heeft)?

Concrete auditvragen daarbij zijn:

- Zijn er duidelijke en gedefinieerde keuzen voor kwaliteitsvaststelling, c.q. blijken van kwaliteit te geven?
- Zijn de standaarden die we aanleggen (bijvoorbeeld voor training van assessoren, afnameprocedures, terugrapportage, advies en dergelijke) van een voldoende niveau (welk niveau)?
- Zijn de uitkomsten van de audit een weergave van de investering die erin gedaan is, met andere woorden: is er een verband tussen inzet en rendement?
- Zijn er eisen gesteld waaraan de assessmentafname en -instrumenten moeten voldoen?

Verder is het belangrijk de aangebrachte articulering uit de vorige stap (welke infrastructuur, welke doelen en dergelijke) te analyseren en te bepalen of de keuzen die gemaakt zijn ook werkelijk kunnen worden getoetst door de concrete criteria te benoemen waaruit kan blijken of de keuzen correct zijn geweest. Dit kan bijvoorbeeld via auditvragen als: is de infrastructuur voldoende werkzaam, worden doelgroepen bediend oftewel is er een instroom?

4. De onderbouwing van het assessment – de conceptuele basis en validering ervan

Het is een grote verantwoordelijkheid assessments af te nemen, want deze kunnen immers leiden tot zwaarwegende beslissingen over ontwikkeling of loopbanen en arbeidsposities van mensen. Er zal daarom een legitimering moeten plaatsvinden van de procedure en de instrumenten om te voorkomen dat een kandidaat beslissingen genomen op grond van het assessment aanvecht. De centrale vraag en zorg is: wat is eigenlijk de kennisbasis waarop het assessment berust?

Deze vraag vertaalt zich direct in de specifiekere vraag: wat is eigenlijk onderzocht en gevalideerd rond de procedure waardoor intern een verantwoording valt te geven van de keuzen die zijn gemaakt en toelichting valt te geven bij de inrichting van met name dit concrete assessment? Het gaat hierbij om de beschikking over een kennisbestand dat aanleiding geeft tot het

verantwoord maken van keuzen en afwegingen rond de concrete inrichting en afname van het assessment-instrumentarium.

Het betreft concrete vragen als:

- Oordelen assessoren betrouwbaar en vanuit een gemeenschappelijk referentiekader?
- Zijn de procedures uitgeschreven en herhaalbaar voor anderen?
- Zijn gegevens verzameld over het gebruik van de assessmentinstrumenten?
- Is er op een afgewogen manier gebruikgemaakt van verschillende assessmentinstrumenten?
- Wordt bijgehouden hoe assessoren tot een afgewogen advies komen?
- Wat is de aard van de feedback die kandidaten ontvangen?

5. Beschikken over een eigen kwaliteitsstandaard/voldoen aan certificeringseisen

Het visieproces en de eigen articulering van assessment (stap 1 en 2) geven aan hoe een instelling op een professionele en eigen wijze opereert. Eigen keuzen rond assessment maken dat men met een uniek product komt en zelfstandig opereert. Toch valt nog aan te geven op welke wijze de eigen keuzen en invullingen overeenkomen met standaarden die extern worden aangelegd over de kwaliteit van assessments. Op de achtergrond speelt de belangrijke vraag: kan onze assessmentaanpak worden geaccrediteerd/gecertificeerd? Op dit moment zijn er (nog) geen landelijke standaarden voor certificering van assessmentprocedures, maar die zullen ongetwijfeld komen. Zodra deze er zijn (via accrediterende organisaties vanuit wetenschappelijk onderzoek of bepaald door landelijke verenigingen van beroepsstandaarden), valt door de instelling zelf te verzekeren hoe de eigen invulling rond assessment spoort met die landelijke standaard.

Concrete vragen die hierbij te formuleren zijn, zijn:

- Welke competenties worden beoordeeld?
- Zijn de competenties gerelateerd aan beroepseisen en werkveldbetrokkenheid?
- Zijn de procedures afgestemd op effectiviteit en kwaliteit in het beroep?
- Hoe verhouden de interne standaarden zich tot standaarden die elders worden gehanteerd?
- Wat zijn de vergelijkende uitkomsten van de procedure met andere beoordelingen?
- Welke externe audit wordt betrokken rond de eigen aanpak?

6. Het auditsysteem onderhouden – afname van de auditinstrumenten

Invoering van assessments vraagt om een voortdurende monitoring of doelen, kwaliteitsstandaarden en

beoogde effecten behaald worden. Om assessment als ‘product’ te kunnen onderhouden wordt van een instelling verlangd (zowel impliciet als steeds meer expliciet) om ‘in de gaten te houden’ wat assessments waard zijn, en wel via audits. We kunnen dit het karakter van de zelftest noemen.

De keuzen voor de eigen missie, de bepaling van standaarden en kwaliteitseisen, dit alles biedt de basis voor een audit als intern controle-instrument. Een audit als zelftest is erop gericht te toetsen of de kwaliteitsstandaarden en de gestelde doelen ook werkelijk worden behaald. Uitvoering van een audit is dan ook uitdrukkelijk een zaak van de betrokken instelling zelf.

Kernvragen daarbij zijn;

- Is er een overall en periodiek evaluatiesysteem in werking om te bepalen wat de kwaliteit van het assessment is?
- Levert het systeem de interne en externe checks voor de beoordeling van de kandidaten?
- Levert de audit operationele data voor onderhoud en verbetering van het assessment?
- Geeft de audit aanwijzingen voor verbetering en vernieuwing van het product?
- Levert de audit een antwoord op over de vraag naar deugdelijkheid?

In kader 1 geven we een illustratie van hoe dit proces is doorlopen bij zeven lerarenopleidingen voortgezet onderwijs, in de daar gangbare assessmentpraktijk.

Audit als meting van kwaliteit

De voorgaande stappen in een audit als proces geven als het ware de grote lijn weer van het verzekeren van kwaliteit als men kijkt naar het vormgeven en organiseren van assessments. Op basis daarvan komt het erop aan een gedetailleerder beoordelingskader te gebruiken voor het bepalen van de kwaliteit van concrete assessmentpraktijken.

Een toetsingskader moet het mogelijk maken direct en op onderdelen aan te geven hoe het gesteld is met de kwaliteit van concrete assessmentpraktijken. Het is goed daarbij een onderscheid te maken tussen enerzijds kwaliteitskenmerken of -indicatoren, dat wil zeggen de inhoudelijke onderdelen van het assessment, en anderzijds de auditcriteria of -standaarden als datgene waarop beoordeeld wordt. Dit is in tabel 1 vervat.

CRITERIA	Aandacht voor	(Uit)werking van	Vertrouwen in
Kenmerken			
Structuur en functie			
Afnameproces			
Meting van prestaties			

Tabel 1.

De keuze voor een toetsingskader oftewel de bepaling van kwaliteitskenmerken en kwaliteitscriteria is essentieel in elke audit. We gaan hier dan ook nader op in.

Een aanzet voor een toetsingskader is verkregen uit een analyse van literatuur binnen de domeinen 'authentic performance assessment' en 'personnel evaluation in public administration', en 'human resource development' en 'teacher appraisal and evaluation'. Aan deze literatuur zijn *criteria* en *kwaliteitskenmerken* ontleend. Een belangrijke bron is verder het (nog summier) onderzoek naar de validiteit van assessments (Lievens, 1998; Darling-Hammond en Snyders, 2000) en de 'Guidelines and Ethical Standards' van de International Task Force on Assessment Centers (2000).

A. Kenmerken

De *kwaliteitskenmerken* zijn gerubriceerd in drie onderwerpen:

- Structuur: de vraag hoe het assessment in elkaar zit, uit welke elementen het bestaat.
- Afnameproces: de vraag op welke manier het assessment verloopt, hoe het assessment wordt gebruikt.
- Meting: hoe komt de vaststelling van competenties of persoonlijke kwaliteiten tot stand?

I. Structuur en functie

De infrastructuurle inbedding van het assessment:

1. Is er sprake van een duidelijk type assessment (intake, ontwikkelingsgericht, doelgroepgericht, domeinafhankelijk); is er een relatie met of inbedding in een natraject (een te volgen training, een ontwikkelingsplan)?
2. Zijn de betrokkenen (kandidaten, assessoren, organisatoren, management, administratie) op hun rol voorbereid en daarvoor toegerust?
3. Is er een duidelijke relatie tussen gemeten competenties en beoogde competenties (welke kennis en vaardigheden van wie worden op welk niveau beoordeeld en met welk doel: geschiktheidsbeoordeling, toelating of plaatsing)?
4. Zijn er uitgewerkte en omschreven procedures voor afname en afhandeling van beoordelingen en zijn de aantallen deelnemers efficiënt en kostendekkend?
5. Is er een koppeling met competentieontwikkeling? Is er aandacht voor implementatie van assessment-

beoordelingen, toelating tot leertrajecten, enzovoort?

6. Als onafhankelijke meting verlangt het assessment een breed draagvlak en erkenning. Spoort de procedure met aanpakken elders, in verband met een mogelijke certificering van de procedure (zie ook hierna validiteit)?

II. Afnameproces

De gebruikte procedures en instrumenten vragen om duidelijke explicitering, en wel ten aanzien van:

1. *De competenties*: dekken deze de startbekwaamheid op een doorzichtige manier? In formulering en wijze van afname van de assessments vinden allerlei impliciete keuzen plaats. Is duidelijk wat waar en met welk instrument gemeten wordt; gaat het om startcompetenties, persoonlijke competenties of beroepskwalificaties?
2. *Het conceptueel kader*: is de afname een eigen invulling en ontwerp of ontleend aan een procedure elders; is de reikwijdte van het instrument vastgesteld of gebruikt men één aanpak voor vele doeleinden en verschillende trajecten?
3. *Beoordelingscriteria*: is duidelijk hoe de criteria bepaald zijn en voor welke competenties; is ook bepaald wat niveaus van beheersing zijn?
4. *Instrumenten*: is verzekerd dat de startcompetenties worden vastgesteld via meerdere assessmentinstrumenten of -methoden? Heeft elk instrument een relatie met andere assessmentvormen? Zijn de instrumenten duidelijk beschreven en gematerialiseerd?
5. *Afnameproces*: is de keuze van de stappen in de afname onderbouwd, hun volgorde, de te meten eigenschappen, inhoudsvaliditeit en de volledigheid ervan?
6. *Observaties*: hebben assessoren een eenduidig beoordelingskader op vaste punten of zijn deze open en te weinig voorgestructureerd, zodat in de beoordeling ook persoonskenmerken op ondoorzichtige wijze meespelen?
7. *Feedback*: wordt er inhoudelijk en persoonsgericht terugkoppeling geboden van de bevindingen?

III. Meting van prestaties

De audit als kwaliteitsborging van het hele assessmentinstrumentarium (inclusief de organisatie eromheen) richt zich verder op vragen rond de wijze van meting, namelijk:

1. Is het assessment representatief voor een bepaald domein/vakgebied; vergt het nog aanzienlijke uitbreiding wil het dekkend zijn voor het beroep? Is het gelegitimeerd in het beroepsveld?
2. Is er zorg voor de standaardisatie en eenduidigheid in beoordeling van diverse competenties?
3. Is de bepaling van beheersingsniveaus en vaststelling van 'cutting points' (geschikt/niet geschikt) onderbouwd?
4. Kunnen assessoren competenties op een vergelijkbare wijze beoordelen, is er een beoordelingskader en zijn de assessoren daarin getraind?
5. Training van assessoren: dit vraagt om aandacht, omdat de inwerking en bekendheid met instrumenten moet worden opgebouwd en onderhouden. Zijn er geen verschillende beoordelingskaders die men hanteert?
6. Is er een duidelijk, inzichtelijk verband tussen de gebruikte observatielijsten en de beoordeling op niveaus/dimensiegedragingen en is bepaald hoe de eindbeoordeling tussen assessoren onderling totstandkomt?
7. Speelt op inzichtelijke wijze de zelfbeoordeling van de kandidaat een rol in de beoordelingsvergelijking met en de uitkomsten van het assessment?
8. Is nagegaan hoe herhaalbaar ('objectief') datgene is wat met het assessment is vastgesteld?
9. Is er verder onderzoek naar de effecten van de procedure? Wat is er geregistreerd en gearhiveerd over de werking van het instrumentarium?

B. Criteria

De doelstelling van de audit is het vaststellen van de kwaliteit van de assessmentpraktijk. De audit moet antwoord geven op vragen als:

- Is er in de structuur van het assessment aandacht voor essentiële kwaliteitskenmerken die gerekend mogen worden tot een goed assessment?
- Is er in de uitvoering van het assessment blijk gegeven van goede kwaliteitshandhaving door middel van gevolgde procedures en werkwijzen?
- Heeft men vertrouwen in de kwaliteit van de *meting* van het assessment, in de zin dat het assessment een goed beeld geeft van zijn/haar competenties?

In de audit wordt telkens informatie verzameld over de individuele kwaliteitskenmerken door deze te beoordelen op de volgende drie *auditcriteria*:

1. aandacht voor ..., dat wil zeggen transparantie en coherentie;
2. werking van ..., dat wil zeggen uitvoerbaarheid en efficiency;
3. vertrouwen in ..., dat wil zeggen validiteit en betrouwbaarheid.

Ad 1

Het is van belang dat het assessment voldoende is toegelicht en beschreven, zowel voor intern gebruik als voor externe verantwoording. Zo wordt verduidelijkt en wordt helder voor betrokkenen welke keuzen zijn gemaakt en hoe concreet en transparant uitwerkingen zijn vormgegeven. Expliciet is aan te geven op welke wijze het assessment is ingericht en herhaalbaar is voor anderen. De vraag die vooropstaat is: is er een toelichting beschikbaar waaruit blijkt dat bewust of op gerichte wijze is gekozen voor de uitwerking van het betreffende kwaliteitskenmerk? Operationeel luidt de vraag: is er informatie beschikbaar over het betreffende kwaliteitskenmerk?

Ad 2

Het is van belang dat de uitvoering van het assessment onder voortdurende kwaliteitszorg staat en mogelijke problemen in de uitvoering met daarvoor geschikte procedures en regels ondervangen worden. De vraag die daarbij vooropstaat is: is zichtbaar in zowel opzet, aanpak als uitvoering van het assessment op welke wijze het assessment wordt uitgevoerd, zodanig dat het voldoet aan gestelde kwaliteitseisen? Operationeel luidt de vraag: is de uitwerking (van het kwaliteitskenmerk) voldoende efficiënt ingericht?

Ad 3

Het is van belang dat het resultaat van het assessment voldoende vertrouwen geniet, ofwel plausibel is, gezien de bewijslast die is verkregen, zodat daarmee een verantwoording valt te geven van de beoordeling die men naar buiten brengt. Het gaat hier duidelijk om een complex, interpretatief oordeel gebaseerd op bewijslast dat openbaar te verantwoorden valt. De vraag die vooropstaat is: is in voldoende mate gewaarborgd dat op deze wijze een valide beeld kan worden gegeven van de kwaliteiten van de kandidaat? Operationeel luidt de vraag: is er vertrouwen in het resultaat?

Het volgende kader bevat als illustratie op de beschreven aanpak in figuur 1 een globaal overzicht van het auditproces zoals afgenomen bij zeven lerarenopleidingen Voortgezet Onderwijs, in de daar gangbare assessmentpraktijk.

DE GEVOLGDE AANPAK IN DE BEPROEVING VAN HET AUDITINSTRUMENT BIJ ZEVEN LERARENOPLEIDINGEN VOORTGEZET ONDERWIJS

Stap 1: Opstart

Doel

In deze fase zijn de benodigde voorbereidingen verricht door met het management en de assessmentcoördinator afspraken te maken over de audit, om de uitvoering soepel en effectief te laten verlopen.

Aanpak

De voorbereiding betrof de volgende activiteiten:

- accordering projectplan en regelen formaliteiten;
- communicatie en uitwisseling van informatie;
- uitnodiging voor deelname aan visiebijeenkomst (stap 2);
- beginnen met verzamelen van inhoudelijke informatie (bronnenonderzoek).

Resultaat

Het resultaat van deze fase is:

- bekendheid met opzet en uitvoering van het auditproces;
- de betrokken partijen in de opleiding zijn geïnformeerd en op de hoogte van deelname;
- concept criteriumlijst is voorgelegd (stap 3);
- organisatie van audit is in de steigers gezet.

Stap 2: Visieontwikkeling

Doel

Het doel van deze stap is in overleg met de opleidingen te komen tot vaststelling van de uitgangspunten van kwaliteitsbeoordeling rond assessment en overeenstemming over het gehanteerde auditinstrument.

Aanpak

Een visiebijeenkomst voor betrokken instellingen en een vervolg daarop zijn georganiseerd. Via een discussie en uitwisseling zijn criteria en kwaliteitskenmerken geformuleerd voor de audit.

Resultaat

Deze stap leidt tot de volgende resultaten:

- bekendheid met de doelstellingen en aanpak;
- steun voor de uitgangspunten van de audit;
- input voor de criteriumlijst die zal worden gebruikt in de audit (hoe meet je de kwaliteit van het assessment?).

Stap 3: Standaardenontwikkeling

Doel

Het ontwikkelen van een lijst met kwaliteitskenmerken is het vertrekpunt voor de audit. Voor het verkrijgen van draagvlak bij de instellingen is het van belang hun ervaring in deze fase erbij te betrekken door hen te laten meedenken over de uiteindelijke lijst van kwaliteitskenmerken.

Aanpak

Bij de vormgeving van het auditinstrument zal een vorm worden gekozen die hanteerbaar blijkt voor de instellingen. Tijdens de audit zal men kennismaken met het instrument.

Resultaat

Het resultaat van deze stap is een instrument in de vorm van een checklist naar kwaliteitskenmerken van assessments, gelieerd aan auditcriteria waarop wordt gemeten.

Stap 4: Gegevensverzameling

Doel

De derde stap betreft de daadwerkelijke auditafname bij de instellingen om uitspraken te doen ten aanzien van de kwaliteit van hun assessmentpraktijk.

Aanpak

De audit bestaat uit de volgende activiteiten:

- de auditoren instrueren over de kwaliteitskenmerken en criteriumlijst;
- visite van de auditoren bij de instelling, bestaande uit:
 - het verkennen van materialen voor het intakeassessment;
 - het voeren van gesprekken met coördinatoren, gebruikers en andere betrokkenen;
 - mogelijk het observeren van de uitvoering van een assessment;
- aanvullende informatie verzamelen via een extra onderzoeksinstrument (vragenlijsten) onder betrokkenen (ex-deelnemers, assessoren, docenten, enzovoort).

Resultaat

De opbrengst van deze fase is:

- analyse van verzamelde gegevens;
- nota met conclusies en aanbevelingen per instelling.

► Stap 5: Terugkoppeling

Doel

Het doel van de terugkoppeling is tweeledig. Ten eerste: het verschaffen van feedback aan de individuele instelling over de resultaten van de audit. Ten tweede: het aanreiken van concrete handvatten waarmee de instelling de kwaliteit van assessments kan blijven meten en verbeteren.

Aanpak

De terugkoppeling naar de instelling vindt plaats in de vorm van een gesprek met betrokkenen van iedere instelling. Een per instelling opgestelde nota vormt hiervoor het vertrekpunt. De terugkoppeling naar de instellingen gezamenlijk vindt plaats door middel van een workshop waarvoor mogelijk geïnteresseerden uit de diverse instellingen worden uitgenodigd.

Resultaat

Na deze stap zijn de volgende resultaten geboekt:

- de instelling heeft een persoonlijke terugkoppeling gekregen op de resultaten van het onderzoek, alsmede handreikingen ter verbetering van de kwaliteit van het intake-assessment;
- het veld is geïnformeerd over de feitelijke kwaliteit van assessmentpraktijken bij de zeven instellingen en heeft aanbevelingen gekregen voor de invoering en verbetering van assessments.

Voor de lerarenopleiding Voortgezet Onderwijs hebben Tillema, Dirkse-Hulscher en Van den Hurk (2002) een auditinstrument ontwikkeld. Het volgende kader bevat een toelichting op het instrument en het gebruik of de afname ervan.

PRAKTIJKVOORBEELD VAN EEN AUDIT-INSTRUMENT VOOR LERARENOPLEIDINGEN VOORTGEZET ONDERWIJS

Het auditinstrument

- Het auditinstrument bestaat uit 460 afzonderlijke *kwaliteitskenmerken* als de meest elementaire bouwstenen waarmee de assessments kunnen worden beoordeeld (zoals de scoringsbetrouwbaarheid tussen assessoren onderling).
- Deze kenmerken zijn gerubriceerd onder verschillende *audittopics* als de onderling samenhangende sets van kenmerken waaruit de audit is opgebouwd. De topics vormen clusters van kenmerken en kunnen onderling weer zijn gegroepeerd tot bepaalde clusters. In de audit zou men zich kunnen beperken tot zo'n selectieve set van kenmerken, dus tot bepaalde topics (bijvoorbeeld een audit naar de training van assessoren).

- De topics zijn uitgewerkt tot vier *auditniveaus*, dat wil zeggen verschillende lagen van detaillering van de topics, lopend van omvattend naar meer specifiek (bijvoorbeeld van de kwaliteit tot de training van assessoren).
- De audit wordt verder gestructureerd door *auditcriteria* die dienen als richtvragen. Dit zijn sturende vragen gesteld vanuit een bepaalde interesse om specifieke informatie over de kwaliteit van het assessment te verkrijgen (bijvoorbeeld: hoe is het gesteld met de beoordeling door assessoren?).

Op het meest algemene auditniveau bestaat het auditinstrument uit de volgende kwaliteitskenmerken c.q. vragen waar de audit zich op richt:

o. Kengetallen

- I. Zijn er cijfers over de afname van het assessment?
- II. Zijn er cijfers bekend over kandidaten?
- III. Zijn er cijfers bekend over assessoren?

A. Doelen en inbedding van het assessment

- I. Is het doel van het assessment helder geformuleerd?
- II. Ligt er een (conceptueel) kader aan ten grondslag?
- III. Is de reikwijdte van het assessment bepaald?
- IV. Verloopt de organisatie van het assessment efficiënt?
- V. Is het assessment ingebed in de opleiding?
- VI. Wordt het assessment verder ontwikkeld/verbeterd?

B. Competenties, standaarden en gedrag

- I. Zijn competenties zorgvuldig vastgesteld?
- II. Dekken de competenties het functioneren in het beroep?
- III. Zijn competenties meetbaar geformuleerd?

C. Instrumentarium

- I. Is het instrumentarium (geheel aan opdrachten) compleet?
- II. Zijn de instrumenten onderbouwd en uitgewerkt?
- III. Bieden instrumenten ruimte voor diversiteit (verschillen)?
- IV. Meten taken/opdrachten/oefeningen relevant gedrag?
- V. Wordt het portfolio effectief gebruikt?
- VI. Is de authentieke opdracht (les) geregeld?
- VII. Worden simulatieoefeningen afgenomen?

D. Assessoren

- I. Worden assessoren zorgvuldig geselecteerd?
- II. Zijn taken en verantwoordelijkheden duidelijk verdeeld?

- III. Worden assessoren getraind/opgeleid?
- IV. Worden assessoren beoordeeld op geschiktheid/kwaliteit?

E. Beoordeling

- I. Is de scoringsprocedure zorgvuldig vastgesteld?
- II. Zijn de condities voor beoordeling optimaal?
- III. Worden competenties op verschillende manieren vastgesteld?
- IV. Vindt scoring op een eenduidige manier plaats?
- V. Is de procedure voor samenstelling van het eindoordeel transparant?
- VI. Zijn eindbeoordeling en eindbeoordelingsgesprek herkenbaar en constructief?

F. Natraject

- I. Worden gegevens gearhiveerd en opnieuw gebruikt?
- II. Wordt er nazorg aan kandidaten verleend?
- III. Is het duidelijk wat na het assessment volgt?

Afname van de audit

De audit als vorm van interne controle is bedoeld om te worden afgenomen als zelftest door een instelling zelf. Men kan daartoe het best een groep (onafhankelijke) auditoren (bijvoorbeeld van buiten de instelling) inschakelen om de audit uit te voeren. Vaststelling van de scoring per kenmerk gebeurt na afloop van een auditbezoek bij een assessmentpraktijk van de instelling op grond van:

- beschikbare bronnen en documenten;
- gesprekken met ter zake betrokkenen: management, coördinatoren, assessoren, kandidaten, secretariaat en andere geïnformeerde personen;
- aanvullende enquêtegegevens over de waardering van de assessmentpraktijk;
- in enkele gevallen directe observatie van de assessmentpraktijk.

De aldus verzamelde gegevens dienen als input voor een afweging per kwaliteitskenmerk tussen de auditoren onderling.

Literatuur

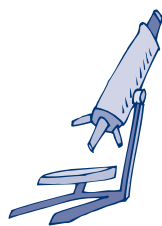
- Bennet, R.E. en W.C. Ward, *Construction vs. choice in cognitive measurement: issues in performance testing and portfolio assessment*. Hillsdale, NJ: Lawrence Erlbaum, 1993.
- Berryman, S.E., 'Learning for the workplace'. *Review of research in education* (19). AERA, Washington, 1993.
- Boot, C. en H.H. Tillema, *Competentiegericht beoordelen in hoger beroepsonderwijs*. Utrecht: Lemma, 2001.
- Broadfoot, P.M., *Education assessment and society*. Milton Keynes: Open University Press, 1996.

- Campbell, D.M., P.B. Cignetti, B.J. Melenyzer, D.H. Nettles en R.M. Wyman Jr., *How to develop a professional portfolio*. Boston: Allyn and Bacon, 1997.
- Darling-Hammond, L. en J. Snyder, 'Authentic assessment of teaching in context'. *Teaching & Teacher Education*, 16, pp. 523-545, 2000.
- Delandshere, G. en A. Petrovsky, 'Assessment of complex performances: Limitations of key measurements assumptions'. *Educational Researcher*, 27 (2), pp. 14-24, 1998.
- Dierik, S., F. Dochy en G. van de Watering, 'Assessment in hoger onderwijs, over implicaties van nieuwe toetsvormen'. *Tijdschrift voor Hoger Onderwijs*, (19) 1, 2001.
- Gipps, C.V., *Beyond Testing*. London: The Falmer Press, 1994.
- Graves, D.H. en B.S. Sunstein, *Portfolio portraits*. Portsmouth: Heinemann Books, 1992.
- Hamilton, M.L., *Reconceptualizing teaching practice; self-study in teacher education*. New York: Corwin Press, 1998.
- Herman, J.L. en L. Winters, 'Portfolio research, a slim collection'. *Educational Leadership*, 52 (2), pp. 48-55, 1994.
- Herriot, P., *Assessment and selection in organizations: methods and practice for recruitment and appraisal*. Chichester: John Wiley, 1989.
- International Task Force on Assessment Centers, 'Guidelines and ethical considerations for assessment center operations'. *Public Personnel Management* (29) 3, pp. 315-331, 2000.
- Jansen, P.W.G. en F. de Jongh, *Assessment centers, een open boek*. Baarn: Spectrum, 1993.
- Klarus, R., *Competenties erkennen*. Academisch proefschrift, Katholieke Universiteit Nijmegen, 1998.
- Lee, D. en D. Beard, *Development centers*. New York: McGraw-Hill, 1994.
- Lievens, F., 'Factors which improve the construct validity of assessment centers'. *International Journal of Selection and Assessment*, 6 (3), pp. 141-152, 1998.
- Olson, M.W., 'Portfolios: education tools'. *Reading Psychology, an international quarterly*, 12, pp. 73-80, 1991.
- Peterson, K.D., *Teacher Evaluation*. Thousand Oaks: Corwin Press, 1995.
- Powell, L.A., 'Realising the value of self-assessment: the influence of the business excellence model on teacher professionalism'. *European Journal of Teacher Education*, 23 (1), pp. 37-49, 2000.
- Redman, W., *Portfolios for Development, a Guide for Trainers and Managers*. London: Kogan Page, 1994.
- Seegers, H.J.J.L., 'Assessment center methode en predictie'. In: A.L.M. Knaapen (red.), *Methoden, technieken en analyses voor personeelsmanagement*. Deventer: Kluwer, 1986.
- Seldin, P., *Successful use of teaching portfolios*. Bolton, Ma.: Andersen Publishing Comp., 1993.
- Shulman, L., E. Haertel en T. Bird, *Toward alternative assessments of teaching, Teacher assessment project*. Report 4-20-88, Stanford University, Stanford, 1988.
- Smith, K., 'Portfolios as an alternative assessment practice in higher education'. In: J.E. Forster (ed.), *University Teaching*. New York: Garland Publishing Inc., 1998.



- Smith, K. en H.H. Tillema, 'Evaluating portfolio use as a learning tool for professionals'. *Scandinavian Journal of Educational Research*, (41) 2, pp. 193-205, 1998.
- Swanson, D.B., G. Norman en R.L. Linn, 'Performance based assessment: lessons from the health profession'. *Educational Researcher*, (24) 5, pp. 5-11, 1995.
- Swieringa, J. en A.F.M. Wierdsma, *Op weg naar een lerende organisatie*, Groningen: Wolters-Noordhoff, 1990.
- Thornton, G.C. en W.C. Byham, *Assessment centers and managerial performance*, New York: Academic Press, 1982.
- Tillema, H.H. (red.), 'Assessment en opleiden in organisaties'. *Capita Selecta* no. 14. Deventer: Kluwer bedrijfswetenschappen, 1993.
- Tillema, H.H., *Development centers; ontwikkelen van competenties*. Deventer: Kluwer bedrijfswetenschappen, 1996.
- Tillema, H.H., 'Design and validity of a portfolio instrument for professional training'. *Studies in Educational Evaluation*, 24(3), pp. 263-278, 1998.
- Tillema, H.H., *Assessment van competenties, van beoordelen naar ontwikkelen*. Alphen a/d Rijn: Samsom, 2001.
- Tillema, H.H., S. Dirkse-Hulscher en V. van den Hurk, 'Naar een audit van assessmentpraktijken in de lerarenopleiding'. *Tijdschrift voor Lerarenopleiders*, 23 (3), pp. 11-20, 2002.
- Topping, K., 'Peer assessment between students in colleges and universities'. *Review of Educational Research*, 68 (3), pp. 249-276, 1998.
- Van der Zee, H., *Opleiden en ontwikkelen*. Deventer: Kluwer bedrijfswetenschappen, 1993.
- Wade, R.C. en D.B. Yarbrough, 'Portfolios; a tool for reflective thinking in teacher education'. *Teaching & Teacher Education*, 12 (1), pp. 63-79, 1996.
- Wiggins, G., 'Teaching to the (authentic) test'. *Educational leadership*, (46) pp. 41-47, 1989.
- Wiggins, G., *Assessing Student Performance, exploring the Purpose and Limits of Testing*. San Francisco: Jossey Bass, 1993.
- Winsor, P., R.L. Butt en H. Reeves, 'Portraying Professional Development in pre-service teacher education'. *Teachers & Teaching*, 5 (1), pp. 59-33, 1999.
- Wood, R., 'Work Samples should be more used (and will be)'. *International Journal of Selection and Assessment*, 2 (3), pp. 166-172, 1994.

$$\alpha^2 + \beta^2 = \gamma^2$$



EXAMENS: SLEUTEL, GEEN SLOT

OP WEG NAAR NIEUWE EXAMENVORMEN IN HET (BEROEPS)ONDERWIJS

ANS GROTENDORST EN SASKIA SCHENNING

3

‘Meer dan nu het geval is zal het gesprek tussen kandidaat en beoordelaar een werkelijk gesprek moeten zijn, waarin vanuit verschillende verantwoordelijkheden gesproken wordt over competenties die wel en niet verworven zijn en die wel of niet de moeite waard zijn om verder te ontwikkelen, hoe dan ook en waar dan ook.’

Dr. Ruud Klarus (2001):
Verzilveren van competenties

INTRODUCTIE

‘Examen’ is een woord dat bij velen van ons herinneringen oproept aan gymlokalen, aula’s of hallen waar tientallen, niet zelden honderden, tafels en stoelen in slagorde staan opgesteld. Waar kandidaten vol spanning afwachten welke vragen uit de verzegelde enveloppen tevoorschijn komen. Waar leerlingen en studenten onder het toezien van surveillanten, binnen een krap bemeten tijdsbestek, de opgaven tot een goede oplossing proberen te brengen.

Examens kunnen altijd rekenen op de belangstelling van de media, vooral als er iets ‘mis’ is of zorgen baart. Regelmatig verschijnen er berichten over (omstreden) civiele effecten van examens, het (gebrek aan) vertrouwen van bedrijfsleven en samenleving in behaalde examenresultaten en over (vermeende) ‘fraudepraktijken’.

In de beeldvorming en berichtgeving over examens voeren klassieke vormen vaak de boventoon: op leerstof- en feitenkennis gerichte toetsing, die tijd- en locatiegebonden is en massaal. Dat beeld hoort nog niet tot het verleden, maar verdient zo langzamerhand wel de nodige nuancering. Want in alle onderwijssectoren is het denkproces over nieuwe examenvormen in volle gang.

De Onderwijsraad¹ heeft in november 2002 een advies uitgebracht aan de Minister en Staatssecretaris van Onderwijs, Cultuur en Wetenschappen (Onderwijsraad, 2002a). Aanleidingen voor de adviesaanvraag waren actuele ontwikkelingen in het onderwijs, de samenleving en het beleid en de gevolgen daarvan voor examinering:

- de wens om meer maatwerk, meer variëteit te leveren;
- de toenemende aandacht voor het denken in competenties (beroepsonderwijs) en leergebieden/vaardigheden (voortgezet onderwijs);
- de spanning tussen centrale kwaliteitsborging – tot uitdrukking komend in landelijke examenprogramma’s en centrale examens – en de autonomie van onderwijsinstellingen;
- de toenemende belangstelling voor alternatieve manieren van examinering en kwaliteitszorg.

In het advies schetst de raad een ontwikkelingsperspectief voor examens, een integrale benadering over

Van:	Naar:
Opleiden	Leren en ontwikkelen
Activiteiten van de docent	Activiteiten van de leerling, de student
(Deel)taken, fragmentatie	Authentieke gehelen, integratie
Behaviouristische kijk op leren	Constructivistische kijk op leren
Leerstof- en vakgericht curriculum	Competentie- en praktijkgericht curriculum

Figuur 1. Veranderend denken over leren en opleiden

sectoren heen, met sectorspecifieke aandacht. De raad wil aanknopingspunten bieden voor de ontwikkeling van een examensystematiek die beter aansluit bij de leerloopbaan van leerlingen en studenten, en die tegelijkertijd recht doet aan maatschappelijke belangen en wensen van de arbeidsmarkt.

LEIDRAAD VOOR DEZE BIJDRAGE

We hebben voor deze bijdrage een selectie gemaakt van enkele thema's uit het advies van de Onderwijsraad, die in het licht van 'deugdelijk assessment' belangwekkend zijn. Deze thema's zijn te beschouwen als dilemma's, spanningsvelden bij de vormgeving van nieuwe examens. In de periode van voorbereiding van het advies is over deze dilemma's van gedachten gewisseld met een groot aantal deskundigen en betrokkenen op het gebied van (nieuwe) examens: onderzoekers en ontwikkelaars, projectleiders en ondersteuners, vertegenwoordigers van inspectie en brancheorganisaties, docenten en opleidingsmanagers.

In dit artikel staan achtereenvolgens drie vragen centraal:

- Wat is de aanleiding tot het veranderend denken over leren, opleiden en examineren?
- Welke functies hebben examens en op welke zou in de komende tijd het accent moeten liggen?
- Welke dilemma's (spanningsvelden) spelen een rol bij de vergelijking tussen klassieke en nieuwere examenvormen en zijn van belang bij het kiezen van nieuwe examenarrangementen?

De titel 'Sleutel, geen slot' geeft de bepleite denkrichting al weer: het accent bij examinering verleggen van selectie en uitsluiting naar talentontwikkeling, en van sluitstuk van onderwijsvernieuwing naar aanknopingspunt voor herontwerp.

VERANDEREND DENKEN OVER LEREN, OPLEIDEN EN BEOORDELEN

Zowel in het voortgezet onderwijs (vmbo, havo, vwo) als in het middelbaar beroepsonderwijs en het hoger onderwijs (hbo en wo) is men op zoek naar alternatieve vormen van opleiden en beoordelen. Deze zoektocht is vooral ingegeven door de noodzaak om een

betere aansluiting te waarborgen tussen onderwijs en maatschappij, onderwijs en arbeidsmarkt en tussen opleidingen onderling. In het beroepsonderwijs manifesteren zich de laatste jaren urgente problemen in branches en beroepen (bijvoorbeeld bij verpleegkundigen en verzorgenden, leraren, politiemedewerkers): grote tekorten op de arbeidsmarkt, de relatie met het werkveld die aan versterking toe is, ontevredenheid over de kwaliteit van de beroepsbeoefenaars, veranderingen in de sector of de maatschappij die een sterk appèl doen op nieuwe bekwaamheden.

In de afgelopen decennia is, mede onder invloed van deze ontwikkelingen, maar ook door voortschrijdende onderwijskundige en leerpsychologische inzichten, het denken over leren en opleiden veranderd. Figuur 1 bevat enkele kenmerken van die verandering (Grotenoordt et al., 2002).

In het beroepsonderwijs vindt een steeds groter deel van het leren plaats in een werksituatie. Dit vraagt om een herziening van de relatie tussen school en arbeidsorganisaties door middel van praktijk- en competentiegericht curricula (zie het volgende kader) en examens die beter passen bij het leren en werken in de praktijk. Ook is er steeds vaker behoefte aan mensen die langs een andere dan de standaard opleidings- en beoordelingsroute geschikt worden verklaard voor de uitoefening van een functie of een beroep.

Het voortgezet onderwijs buigt zich over de vraag hoe de vernieuwing in de onderwijsprogramma's (vmbo, tweede fase) versterkt kan worden door nieuwe examenvormen in plaats van geblokkeerd door traditionele. Het hoger onderwijs 'worstelt' met de waarde van het vwo-diploma, de invoering van de bachelor-masterstructuur en met haar eigen examenpraktijken.

Examenproblemen worden meestal per schoolsoort of per opleiding opgelost. Deze 'verkokerde' aanpak gaat wringen. Schoolsoorten onderling en scholen en branches houden er ieder hun eigen opvattingen en werkwijzen op na, spreken verschillende 'talen', verstaan elkaar niet. Er is steeds meer behoefte aan een kader voor de ontwikkeling van examens voor het onderwijs als geheel. Een benadering die de functie en techniek van examens plaatst in het perspectief van de *leerloopbaan* van de deelnemers aan dat onderwijs.

COMPETENTIEGERICHT ONDERWIJS

De term 'competentie' (bekwaamheid) wordt in het onderwijs vaak gehanteerd zonder dat duidelijk is wat er precies mee wordt bedoeld. Dat bemoeilijkt de communicatie over competentiegericht opleiden en beoordelen. Een veel gebruikt synoniem van competentie is (handelings)bekwaamheid. Van het begrip 'competentie' circuleren nogal wat definities. Thijssen en Lankhuijzen (2000, p. 133) geven een algemene typering, die de essentie van veel andere beschrijvingen omvat: 'Een cluster van kennis-, vaardigheids- en attitude-elementen, geschikt om met een geheel van als kritisch beschouwde probleemsituaties in de praktijk adequaat om te gaan.'

De Onderwijsraad heeft tegelijk met het advies 'Examinering in ontwikkeling' een studie gepubliceerd naar het competentiebegrip (Onderwijsraad, 2002b). De studie gaat ervan uit dat volledige consensus over de inhoud van het begrip niet realiseerbaar is, maar dat het competentiebegrip in ieder geval drie noodzakelijke dimensies dient te bevatten: specificiteit, integrativiteit en duurzaamheid. Niet noodzakelijk maar wel relevant (voorwaardelijk voor de verwerving van competenties) zijn de dimensies 'handelingsgerichtheid', 'leerbaarheid' en 'onderlinge afhankelijkheid'.

Noodzakelijke dimensies:

- *Specificiteit* Het verwerven en toepassen van competenties vindt plaats in contexten. Sommige competenties zijn min of meer generiek (bijvoorbeeld leercompetenties), voor andere competenties (bijvoorbeeld beroepsgerelateerde) geldt dat minder.
- *Integrativiteit*. Een competentie is een samenhangend geheel van elementen dat nodig is voor probleemoplossend handelen. De elementen hebben betrekking op kennis, vaardigheden, houdingen en eigenschappen.
- *Duurzaamheid*. Kenmerkend voor een competentie is dat er sprake is van een zekere duurzaamheid. De inhoud van een competentie is niet onveranderlijk, maar kan wel 'een tijdje mee'.

Relevante dimensies:

- *Handelingsgerichtheid*. Voor de ontwikkeling en toetsing van een competentie is probleemoplossend handelen voorwaardelijk. Competenties ontwikkelen zich niet in het luchtledige, maar in concrete handelingsituaties. Handelen omvat ook mentale activiteiten die niet waarneembaar zijn, maar wel – ten behoeve van de toetsing – tot op zekere hoogte inzichtelijk gemaakt kunnen worden.
- *Leerbaarheid*. Het vergt inspanning en tijd om een competentie te ontwikkelen. De leerbaarheid van een competentie varieert al naar gelang (stabiele, moeilijk veranderbare) persoonskenmerken essen-

tieel worden geacht voor het beheersen ervan. Verder wordt de leerbaarheid beïnvloed door iemands leervermogen.

- *Onderlinge afhankelijkheid*. Competenties staan in een bepaalde relatie tot elkaar. Meestal is er sprake van afhankelijkheden die niet voorwaardelijk zijn, maar wel relevant voor onderwijsdoeleinden.

FUNCTIES VAN EXAMENS

De functies die examens hebben, lopen uiteen van het aangeven van een individueel eindniveau tot het monitoren van het landelijke onderwijspeil.

De Onderwijsraad (2002a) onderscheidt een kwalificerende functie, een paspoortfunctie, een kwaliteitsfunctie en een maatschappelijke functie van examens:

1. Bij de *kwalificerende functie* gaat het om het vaststellen van de voortgang van een leerling of student in een leertraject, of om het vaststellen van het eindniveau dat hij of zij heeft behaald na het doorlopen van een opleiding.
2. De *paspoortfunctie* verwijst naar:
 - toelating of uitsluiting (een student krijgt wel of geen toestemming de opleiding voort te zetten, bijvoorbeeld na de propedeuse in het hoger onderwijs);
 - allocatie (een student krijgt een passende plaats in een vervolgopleiding of op de arbeidsmarkt);
 - selectie (een vervolgopleiding of een actor op de arbeidsmarkt kiest een student uit op grond van behaalde prestaties).
3. De *kwaliteitsfunctie* duidt erop dat examens:
 - een sturende werking (kunnen) hebben op inhouden van programma's en handelwijzen van studenten en docenten;
 - studenten (kunnen) motiveren tot leren;
 - de talentontwikkeling van studenten (kunnen) prikkelen en stimuleren;
 - informatie (kunnen) bieden aan ouders, prestaties (kunnen) verantwoorden aan derden;
 - de communicatie tussen opleidingen en het afnemend veld en tussen docenten onderling (kunnen) bevorderen.
4. De *maatschappelijke functie* betreft de rol die examens hebben bij:
 - het monitoren en beïnvloeden van het algemene onderwijspeil;
 - het beoordelen van prestaties van het onderwijs in relatie tot de kosten;
 - het bijdragen aan nationale doelstellingen als gelijkheid van kansen en nationale identiteit;
 - onderwijsinnovatie: examens vormen dan niet alleen het sluitstuk van de uitvoering van het onderwijsprogramma, maar zijn ook beginpunt van het curriculumontwerp;

- ‘benchmarking’ in internationaal opzicht (diplomavergelijking).

Examens vervullen vaak meerdere functies tegelijk. De kwalificerende functie van examens is echter leidend. In dit verband spreekt men van het ‘civiel effect’ van examens, dat wil zeggen: het vertrouwen dat de samenleving, vervolgoopleidingen en het bedrijfsleven hebben in de resultaten die kandidaten op een examen hebben behaald. Het waarborgen van het civiel effect van examens is een bescherming voor de student tegen ondeugdelijk onderwijs. Deze waarborg is krachtiger naarmate de herkenbaarheid en transparantie van examens groter zijn.

DILEMMA’S

Bij de keuze voor nieuwe examenvormen en -arrangementen spelen dilemma’s, spanningsvelden een rol. Verschillende van deze dilemma’s zijn ook op te vatten als *ontwerpvragen*. De Onderwijsraad heeft een aantal dilemma’s geïdentificeerd en deze in clusters geordend. Figuur 2 bevat een overzicht; daarna volgt een uitwerking van de vragen afzonderlijk.

1. Functies

- Selectief of adaptief?
- Leerproces of leerresultaat?

2. Verantwoordelijkheidsverdeling

- Formele en feitelijke verantwoordelijkheid?
- Contextgebonden of gedecontextualiseerd?

3. Positionering

- Leerwegaafhankelijk of leerwegafhankelijk?
- Onderwijsuitvoering en -beoordeling gescheiden?
- Retrospectief of prospectief?
- Standaardisatie of destandaardisatie?

4. Inhoud en afname

- Vak- en leerstofgericht of competentiegericht?
- Absolverend of integratief?

5. Conditie voor nieuwe examenvormen

- Organisatie
- Financiering
- Implementatie

Figuur 2. Dilemma’s nieuwe examens (ontwerpvragen)

Functies

De kwalificerende functie is de belangrijkste, primaire functie van examens. Andere functies lijken de laatste tijd echter aan belang te winnen. Deze staan

niet zelden op gespannen voet met de kwalificerende functie.

Selectief of adaptief?

Er bestaat spanning tussen de selectieve en adaptieve rol van examens. Bij adaptieve examens (adaptie = aanpassing) ligt het accent op de persoonlijke ontwikkeling van de leerling, bij selectieve examens op toelating of uitsluiting van kandidaten voor vervolgoopleiding of beroep. In het vmbo en op de niveaus 1 en 2 van het middelbaar beroepsonderwijs zijn het voorkomen van uitval en het bevorderen van doorstroming en ‘employability’ van cruciaal belang. Leerlingen zijn daar sterk gebaat bij een adaptieve rol van examens, bij ontwikkelingsgericht toetsen, bij ondersteuning van hun eigen leerloopbaan. In het hoger onderwijs klinkt de roep om sterkere selectie van studenten door invoering van toelatingstoetsen (Begeleidingscommissie Decentrale Toelating, 2002). Toelatingstoetsen (voor vervolgoopleidingen of branches) zijn te begrijpen in het licht van de paspoortfunctie van examens, maar verhouden zich moeizaam tot principes als toegankelijkheid en talentontwikkeling (kwaliteitsfunctie). Ze druisen ook in tegen het actuele streven naar ‘doorlopende leerlijnen’.

De vraag wat ‘deugdelijk assessment’ is, zal hier per actor tot een verschillend antwoord leiden. Met examens zijn immers diverse, soms tegengestelde belangen gemoeid.

Leerproces of leerresultaat?

Hoewel het geen primaire functie is van examens, vormen zij een essentieel aanknopingspunt voor verbetering en vernieuwing van het onderwijs. Examens zijn dan onderdeel van het (her)ontwerpproces van een opleiding; in de uitvoering blijven zij een afsluitende functie houden. Ingrijpende onderwijsvernieuwingen zijn niet zonder veranderingen in het examensysteem te realiseren (zie het artikel van Bolhuis en Van Royen in dit katern en zie Grotendorst et al., 2002).

Verantwoordelijkheidsverdeling

De verantwoordelijkheid voor inhoud en procedure van de examens is per onderwijssector wettelijk verschillend geregeld.

Formele en feitelijke verantwoordelijkheid?

Bij nieuwe examenvormen is het nodig de verantwoordelijkheidsverdeling opnieuw te doordenken, afhankelijk van de gemaakte keuzes. Welke ruimte biedt de wetgeving voor veranderingen in het examensysteem? Waar is aanpassing van wet- en regelgeving gewenst? In hoeverre moeten externe instanties betrokken zijn bij het ontwerp van exameninhouden en -procedures? Hoe moet verantwoording over exa-

menresultaten worden afgelegd, door wie, aan wie? Verder dienen zich vragen aan op het gebied van validering van examens, kwaliteitsbewaking en toezicht. Wie is verantwoordelijk voor het leren en beoordelen op de werkplek? Welke verantwoordelijkheid hebben docenten, de inspectie? En welke verantwoordelijkheid heeft de student die zijn eigen portfolio beheert? (Bom, Klarus en Nieskens, 1997). Bij de beantwoording van deze vragen moet, volgens de Onderwijsraad, het civiel effect (dus de kwalificerende functie) van nieuwe examenvormen steeds leidend zijn.

Contextgebonden of gedecontextualiseerd?

Een ander dilemma binnen dit cluster is het contextualiseren of juist decontextualiseren van opleidingen en examens. Deze vraag is vooral in het beroepsonderwijs actueel. Daar stelt men de gewenste eindresultaten van

Competenties, zeker beroeps- en praktijkgerelateerde, zijn *per definitie* contextgebonden

een opleiding vast in overleg met de branche, met beroepsgroepen of in overleg met werkveldcommissies. Betrokkenheid van de praktijk bij het opstellen van beroeps- en competentieprofielen en bij beoordelingen vergroot de acceptatie van beoordelingsresultaten (dus het civiel effect). Het middelbaar beroepsonderwijs staat op dit moment voor een wending naar een competentiegerichte kwalificatiestructuur (ACOA, 1999; Van der Kolk en Willemse, 2002). In dat kader wordt nu voorgesteld eindtermen (competenties) te formuleren op een relatief hoog abstractieniveau en tamelijk los van de context, de branche of de beroepsgroep (Colo, 2002). Dit is een potentiële, serieuze bedreiging van het civiel effect van opleidingen en examens. Competenties, zeker beroeps- en praktijkgerelateerde, zijn *per definitie* contextgebonden.

Klarus (2001, p. 21) verwoordt dit als volgt:

- Competenties op zichzelf zijn theoretische, abstracte constructies. Hun verschijningsvorm is altijd gecontextualiseerd. Ze zijn ook alleen in hun contextualiteit, in gesitueerde vorm te beoordelen.
- Geoperationaliseerde, reëel bestaande en waarneembare competenties nemen de vorm aan van een volledig handelingsproces: voorbereiden/plannen, uitvoeren, evalueren/controleren. Deze handelingsvolgorde zou ook ten grondslag moeten liggen aan competentiebeoordeling.
- Kennis en inzicht kunnen niet op zichzelf, maar alleen in geëxterioriseerde (letterlijk: naar buiten verplaatste) vorm beoordeeld worden. Omdat kennis en inzicht als onderdeel van competenties sterk

gecontextualiseerd zijn, staat het toepassen van kennis en inzicht in de beoordeling centraal.

Positionering

Bij de vraag hoe nieuwe examenvormen te positioneren, spelen verschillende dilemma's een rol.

Leerwegaafhankelijk of leerwegafhankelijk?

Examens kunnen leerwegafhankelijk of leerwegaafhankelijk zijn. Een geheel leerwegaafhankelijk systeem vertegenwoordigt de meest vergaande scheiding tussen begeleiding van het leerproces en beoordeling van de leerresultaten van een kandidaat. Staatsexamens en EVC-procedures² zijn hiervan voorbeelden. Onderwijs (leerproces, leeromgevingen) en examens (beoordeling) horen een sterke samenhang te vertonen. Anders gezegd: het curriculum moet intern en extern consistent zijn.³ Voor elke opleiding (sector) is het van belang een balans te vinden tussen onderwijsinhoud en examens en een mate van onafhankelijkheid die passend is.

Onderwijsuitvoering en -beoordeling gescheiden?

De toenemende autonomie van scholen maakt de noodzaak tot onafhankelijke kwaliteitsborging van examens groter. Bij schoolgebonden examens of instuutsgebonden tentamens staan twee opvattingen tegenover elkaar: onderwijsuitvoering en -beoordeling in één hand, of uitvoering en beoordeling gescheiden. Om het civiel effect van opleidingen en examens te waarborgen, is een zekere mate van onafhankelijkheid bij de beoordeling wenselijk. Daar is echter geen geheel ontkoppeld opleidings- en examensysteem voor nodig. In veel gevallen kan het toereikend zijn om de docent die begeleidt niet als enige ook verantwoordelijk te stellen voor de beoordeling van zijn of haar student. (Dit geldt niet wanneer het sterk in het belang van de leerling is dat leerinspanningen worden meegewogen, zoals in het vmbo en bij de niveaus 1 en 2 van het mbo).

Retrospectief of prospectief?

Examens kunnen retrospectief zijn ('terugkijkend'), waarbij de nadruk ligt op het 'overhoren', de reproductie van hetgeen de leerling in de opleiding heeft geleerd.

Examens kunnen ook prospectief ('vooruitkijkend', 'aankondigend') zijn. In dat geval is een examen een voorspeller van het functioneren van een kandidaat in de maatschappij, van zijn of haar studiesucces of functioneren in de beroepspraktijk. De voorspellende waarde van examens krijgt de laatste tijd meer aandacht, lijkt aan belang te winnen. Dat werpt ook een nieuw licht op de kwalificerende functie en paspoortfunctie van examens.

Standaardisatie of destandaardisatie?

De onderwijssectoren laten grote verschillen zien in de mate waarin examens zijn gestandaardiseerd. In het mbo gaat men de komende tijd op landelijk niveau vorm geven aan nieuwe onderwijs- en beoordelingsconcepten. In het hbo ligt het initiatief doorgaans bij onderwijsinstellingen, al of niet in overleg met het regionale werkveld. Het voortgezet onderwijs is de enige sector die naast een schoolexamen ook een centraal examen kent. In het wetenschappelijk onderwijs is van standaardisatie vrijwel geen sprake.

Mét een toenemende flexibilisering van het onderwijs en variatie aan leerwegen groeit de behoefte aan standaarden om de kwaliteit van het onderwijs aan af te meten. Enige mate van standaardisering lijkt gewenst met het oog op de maatschappelijke erkenning van examens en om ongewenste sociale effecten van examens tegen te gaan (bijvoorbeeld meer selectie). Het is beslist niet de bedoeling dat externe standaarden leiden tot starre, centralistische opleidings- en beoordelingsconcepten. Zowel landelijke als regionale kaders zijn mogelijk. Van belang is dat onderwijsinstellingen gestimuleerd worden transparant te zijn over de leerresultaten van hun deelnemers en over toegekende diploma's.

Inhoud en afname (vorm)

Ontwikkelingen in de maatschappij en in het bedrijfsleven én veranderende opvattingen over leren en opleiden (zie de paragraaf) leiden tot andere eisen aan vorm en inhoud van onderwijs en examens.

Vakgerichte kennis of competenties?

De overgang van een op leerstof en feitenkennis gericht schoolsysteem naar competentie- en praktijkgerichte leertrajecten heeft directe gevolgen voor de manier waarop beoordeeld wordt.

Het is daarbij de vraag of we moeten verlangen naar een totale 'omslag' naar competentiegerichte onderwijs- en examenvormen. Het antwoord op die vraag verschilt per onderwijssector. Competentiegericht opleiden en beoordelen zijn 'in' en publicaties erover hebben voornamelijk een hoog 'hallelujagehalte'. Competentiegericht opleiden houdt ontegenzeggelijk beloftes in, maar is geen panacee voor alle problemen die zich in een onderwijssector voordoen. Algemeen gevoeld is het verlangen opleidingen beter op elkaar af te stemmen, binnen en tussen sectoren (bijvoorbeeld in de 'beroepskolom' vmbo-mbo-hbo). Competentiegericht opleiden kan een middel zijn om een betere afstemming te realiseren (zie de paragraaf 'Veranderend denken over leren, opleiden en beoordelen'). Wanneer beroepsopleidingen van verschillende niveaus zich bedienen van dezelfde ontwerpprincipes en taal, kan dat bevorderend werken op een soepele doorstroming.

Wat het toetsen van competenties betreft: zie daarvoor ook wat is opgemerkt bij 'contextgebonden of gedecontextualiseerd'.

Absolverend of integratief?

Een ander dilemma/ontwerpvoorbeeld dat de Onderwijsraad bij de inhoudsdimensie van examens heeft benoemd, is het onderscheid tussen absolverend en integratief beoordelen. Absolverend beoordelen (letterlijk: steeds een deel van het examen kwijtschelden)

De vraag 'Absolverend of integratief beoordelen, wat is deugdelijk?' is vooral relevant in beroepsgerichte opleidingen

vindt plaats in de loop van de opleiding op afzonderlijke onderdelen. Een voorbeeld zijn de deelkwalificaties in het middelbaar beroepsonderwijs. Integratief beoordelen betreft samenhangende gehelen van elementen (kennis, vaardigheden, houdingen, eigenschappen) noodzakelijk voor probleemoplossend handelen.

De vraag 'Absolverend of integratief beoordelen, wat is deugdelijk?' doet zich in alle onderwijssectoren voor, maar het sterkst in beroepsgerichte opleidingen. Deze vraag is ook van toepassing op EVC-procedures, die vaak een absolverend karakter hebben.

Bij alle opleidingen zou naast relatief losse, absolverende beoordelingen altijd sprake moeten zijn van integratieve toetsing van bekwaamheden. Dat zou het huidige, veelal fragmentarische, toetsen terugdringen. Van belang bij integratieve (competentie)beoordelingen is dat deze vooraf bij de kandidaat bekend zijn, zodat een gerichte voorbereiding mogelijk is (zie het artikel van Bolhuis en Van Royen in dit katern). Integratieve beoordelingen kunnen de kans op overmatige 'near transfer' reduceren. Een student moet in haar of zijn vervolgopleiding of beroep immers in staat zijn tot transfer van vaardigheden naar verschillende contexten.

Conditie voor nieuwe examenvormen

De Onderwijsraad heeft behalve dilemma's enkele conditionele kwesties benoemd die de inhoud en vormgeving van examens beïnvloeden en die bij alle bovenstaande vragen en spanningsvelden een rol spelen:

- de organisatie van nieuwe examenvormen;
- de financiering van nieuwe examenvormen;
- de implementatie van nieuwe examenvormen.

Organisatie

Nieuwe examenvormen hebben gevolgen voor de organisatie van opleidingstrajecten en de relaties die

onderwijsinstellingen met hun omgeving aangaan. Meer flexibiliteit in vorm, afnametijdstip en afnamelocatie van examens leidt tot een andere, complexere organisatie. Nieuwe exameninhouden en -vormen brengen veranderingen teweeg in de wijze waarop docenten hun onderwijs inrichten. Niet alleen investeringen in inhoud en instrumentatie zijn van belang, maar ook een verdere ontwikkeling van de deskundigheid van docenten, coaches en praktijkbeoordelaars. Bij nieuwe examenvormen speelt de externe consistentie en validiteit een steeds grotere rol. Dat vraagt een hoge mate van betrokkenheid van het afnemend veld en de bekwaamheid om die te kunnen organiseren.

Financiering

De vraag hoe de nieuwe examenvormen en -arrangementen te financieren, duikt regelmatig op. Assessments en nieuwe examenvormen staan te boek als arbeids- en tijdsintensief, dus duur. Het is van veel belang om geschikte financieringsmethoden te vinden of te ontwikkelen. Meer variatie betekent een flexibeler bekostigingssysteem. Hoe blijft een flexibele financieringsmethode controleerbaar? Welke mate van standaardisering is nodig om de kosten te beheersen? Hoe kan een evenwicht worden gevonden tussen variatie en standaardisatie? Het verdient aanbeveling om financiële consequenties van nieuwe examenvormen in samenhang te bekijken met het opleidingssysteem als geheel en de inrichting van het onderwijs. Nieuwe examenvormen kunnen leertrajecten efficiënter en vaak korter maken. De huidige praktijk met zijn klassieke kengetallen en sturingsmechanismen (roosters, teldata) werpt hiervoor nog veel belemmeringen op. Het is mogelijk kosten te besparen bij een gelijktijdige kwaliteitsverbetering van onderwijs en examens.

Implementatie

De implementatie van nieuwe examenvormen vraagt om gunstige voorwaarden. In de ene sector of branche zal dit gemakkelijker gaan dan in de andere. Een zorgvuldige voorbereiding, bijvoorbeeld met pilots of experimenten, verdient aanbeveling. Het is daarom van belang dat nieuwe examenvormen deel uitmaken van het langetermijnbeleid van instellingen en overheid.

RÉSUMÉ

Het denken over leren, opleiden en beoordelen (examineren) is de afgelopen jaren flink in beweging gezet. Bij de introductie en verdere ontwikkeling van nieuwe examenvormen en -arrangementen is het goed de functies van examens opnieuw te doordenken. Van belang lijkt een herdefiniëring van civiele effecten. Wat betekenen veranderingen voor de perceptie van examens door vervolgopleiding, arbeidsmarkt en maatschappij?

De Onderwijsraad beveelt een integrale benadering aan van examens, binnen en tussen onderwijssectoren. De ontwikkeling van nieuwe beoordelingsvormen komt onvoldoende tot zijn recht wanneer uitsluitend een sectorspecifieke invalshoek wordt gekozen. Het gaat om 'een versterking van de steun die een geïntegreerd geheel van verrijkte examenvormen kan geven aan de leerloopbaan, de leerbiografie van mensen tijdens hun jonge jaren en de levensperiodes die daarop volgen'. In dit perspectief zijn examens een sleutel naar de toekomst.



Literatuur

- Adviescommissie Onderwijs-Arbeidsmarkt (ACOA), *Een wending naar kerncompetenties. De betekenis van kerncompetenties voor de versterking van de kwalificatiestructuur secundair beroepsonderwijs*. Den Bosch: ACOA, 1999.
- Begeleidingscommissie Decentrale Toelating, *Decentrale toelating in een veranderende context*. (Analysedocument behorende bij tussenrapportage decentrale toelating 2001/2002.) 2002.
- Bom, W., R. Klarus en M. Nieskens, *Portfolio in opleiding en bedrijf*. 's-Hertogenbosch: Cinop, 1997.
- Colo, *Samen werken aan leren. Naar een competentiegerichte kwalificatiestructuur voor het middelbaar beroepsonderwijs*. Zoetermeer: Colo, 2002.
- Grotendorst, A., 'Non scholae, sed vitae discimus. Een competentiegerichte curriculumbenadering'. In: M. Rondeel en S. Wagenaar (red.), *Kennis maken: leren in gezelschap* (pp. 193-215). Schiedam: Scriptum Management, 2002.
- Grotendorst, A., M. Jellema, I. Stam, M. van der Vegt en C. Zandbergen, *Leren in veiligheid. Het nieuwe politieonderwijs in maatschappelijk perspectief*. Apeldoorn: LSOP Politie Onderwijs- en Kenniscentrum, 2002.
- Kessels, J.W.M., *Succesvol ontwerpen. Curriculumconsistentie in opleidingen*. Deventer: Kluwer, 1996.
- Klarus, R., *Verzilveren van competenties*. Wageningen: Stoas Onderzoek, 2001.
- Kolk, J. van der en P. Willemse, *Van eindtermen naar competentiebeoordeling. Vernieuwing van examinering in de BVE-sector*. 's-Hertogenbosch: Cinop, 2002.
- Onderwijsraad, *Examinering in ontwikkeling. Een ontwikkelingsperspectief voor examens in het voortgezet onderwijs, middelbaar beroepsonderwijs en hoger onderwijs*. Den Haag: Onderwijsraad, 2002a.
- Onderwijsraad, *Competenties: van complicaties tot compromis. Over schuifjes en begrenzers*. (Studie in opdracht van de Onderwijsraad uitgevoerd door prof. dr. Jeroen J.G. van Merriënboer, dr. Marcel R. van der Klink en drs. Maaïke Hendriks; Onderwijstechnologisch Expertisecentrum, Open Universiteit Nederland). Den Haag: Onderwijsraad, 2002b.
- Thijssen, J.G.L. en E.S.K. Lankhuijzen, 'Competentiemanagement en employabilitystrategie: een kader voor beleidskeuzes'. In: F. Glastra en F. Meijers (red.), *Een leven lang leren. Competentieontwikkeling in de informatiesamenleving*. 's-Gravenhage: Elsevier bedrijfsinformatie bv, 2000.

Noten

1. De Onderwijsraad is een onafhankelijk adviescollege, dat adviseert over de hoofdlijnen van het beleid en de wetgeving op het gebied van onderwijs. De raad adviseert (gevraagd en ongevraagd) de ministers van Onderwijs, Cultuur en Wetenschappen (OCenW) en van Landbouw, Natuurbeheer en Visserij. Ook de Eerste en Tweede Kamer der Staten-Generaal kunnen om advies vragen.
2. Erkenning van (hoe dan ook) verworven competenties.
3. Zie Kessels, *Succesvol ontwerpen. Curriculumconsistentie in opleidingen* (1996). Interne consistentie: logische samenhang tussen curriculumonderdelen. Externe consistentie: gedeelde opvatting tussen docenten onderling en docenten en praktijkveld over de gewenste resultaten van het onderwijs.



• LSOP

• Centraal Examenbureau Politie

Politie Onderwijs- en Kenniscentrum

COMPETENT BLAUW OP STRAAT

PROEVEN VAN BEKWAAMHEID IN HET VERNIEUWDE POLITIEONDERWIJS

BART BOLHUIS EN RICHARD VAN ROYEN

4

Het LSOP Politie Onderwijs- en Kenniscentrum is in 2002 gestart met vernieuwd politieonderwijs, bekend onder de naam 'Politieonderwijs 2002'. Examineren gericht op het meten van kennis en inzicht (bijvoorbeeld door middel van een meerkeuzetoets) is grotendeels vervangen door het beoordelen van bekwaamheden (competenties) van politiestudenten door middel van verschillende examenopdrachten. De ontwikkeling, afname en beoordeling van examens vindt plaats onder de verantwoordelijkheid van het Centraal Examenbureau Politie (CEP).

In dit artikel beschrijven wij, toetskundigen van het CEP, hoe wij zorgdragen voor 'deugdelijk assessment' door middel van de proeven van bekwaamheid. Eerst gaan we in op het politieonderwijs en de plaats van de proeve van bekwaamheid in het ontwerp van het competentiegericht curriculum. Vervolgens richten we ons op de maatregelen die we nemen om de kwaliteit van onze proeven van bekwaamheid te garanderen. Hierbij maken we onderscheid tussen de ontwikkeling en de uitvoering van de proeven. We illustreren dit aan de hand van concrete voorbeelden. Dit artikel sluiten we af met enkele leerervaringen waar men rekening mee kan houden bij de ontwikkeling van proeven van bekwaamheid.

Vanessa Koolmees (student allround politiemedewerker in de politieregio Utrecht): 'De proeve van bekwaamheid is een examen in een echte, gewone en natuurlijke situatie. Het was wel wennen om met een examinerator van het CEP en mijn eigen praktijkcoach de straat op te gaan om hen te laten zien dat ik wel competent ben. Vooral bij de proeve 4.1.1 "Toezicht houden in publiek domein" liepen nogal wat uniformen mee; je zag soms de mensen vanachter de geraniums de gordijnen opzij schuiven en medelijdend naar je kijken. Je voelde de ogen van de examinerator in je rug branden.'

Erik van Herk (student politiemedewerker bij het Korps Landelijke Politie Diensten): 'Ik vind het veel beter dat de proeven op straat worden afgenomen in de reële werkomgeving en niet nagespeeld op het politieschoolplein.'

COMPETENTIEGERICHT POLITIEONDERWIJS

Politiefunctionarissen oefenen hun beroep uit midden in een maatschappij die continu in beweging is. Veiligheid en leefbaarheid zijn belangrijke thema's waarbij de politie een cruciale rol speelt. Dit vraagt om competente politiefunctionarissen. Politieonderwijs 2002

is hierop een antwoord waarbij competentiegerichte en duale leertrajecten sleutelbegrippen zijn.

Competentiegericht

In tegenstelling tot het traditionele onderwijs, waarin het curriculum is ingedeeld in vakken met eigen doelstellingen die omschreven zijn in kennis, vaardigheden en attitude, is het vernieuwde politieonderwijs gericht op verwerving van competenties. De Adviescommissie Onderwijs-Arbeidsmarkt (ACOA, 1999) definieert een competentie als ‘het vermogen van een persoon om in beroeps-, (leer- en maatschappelijke) situaties op adequate wijze proces en productgericht te handelen’. Hierdoor zijn de bekwaamheidseisen uit de beroepspraktijk bepalend voor curriculumontwikkeling en competentiebeoordeling.

Duaal leren

Naast competentiegericht is het vernieuwde politieonderwijs duaal. Dit is een voorwaarde voor succesvol competentiegericht onderwijs (Grotendorst et al., 2002). De student combineert het leren op een van de instituten van het LSOP met het werkend leren in een van de politiekorpsen. De student heeft een werkgever-werknemerrelatie bij een korps en is opsporingsbevoegd.

De instituten en korpsen zijn verantwoordelijk voor de kwaliteit van de leeromgevingen, waarin de student de mogelijkheid krijgt om de benodigde competenties te verwerven. Tijdens de opleiding neemt de bekwaamheid van de student toe door het behalen van de proeven van bekwaamheid. Hierdoor breidt de student bevoegdheden uit en is in toenemende mate inzetbaar in het korps (LSOP, 2001).

ONTWERPMETHODIEK

Bij het ontwerpen van de nieuwe politieopleidingen wordt de ontwerpmethodiek ‘competentiegericht curriculum’ gebruikt (Grotendorst et al., 2002; Kessels, 1996). Deze ontwerpmethodiek kent een aantal fasen. In het kader van dit themanummer ligt het accent in dit artikel op fase 3: de proeven van bekwaamheid¹. Voor een goed begrip is het echter van belang om aandacht te schenken aan de relatie met de voorgaande fasen.

Fase 1: Beroepsprofiel

De eerste fase van de ontwerpmethodiek richt zich op het vaststellen van beroepsprofielen. Dit is het profiel van een beginnend beroepsbeoefenaar en geldt als gemeenschappelijk startpunt voor volgende fasen in de ontwerpmethodiek. Voor het initiële politieonderwijs hebben de Minister van Binnenlandse Zaken en

Koninkrijksrelaties en de Minister van Justitie vijf beroepsprofielen vastgesteld. Deze komen overeen met de niveaus in de reguliere kwalificatiestructuur²: assistent politiemedewerker (niveau 2 mbo), politiemedewerker (niveau 3 mbo), allround politiemedewerker (niveau 4 mbo), politiekundige bachelor (hbo), politiekundige master (hbo-master). Vervolgens zijn elf specialistische beroepsprofielen vastgesteld voor het postinitiële politieonderwijs.

Fase 2: Kernopgaven

De tweede fase van de ontwerpmethodiek betreft de ontwikkeling van kernopgaven, op basis van de beroepsprofielen. De ACOA definieert kernopgaven als: ‘De centrale opgaven en problemen waarmee een beroepsbeoefenaar regelmatig in aanraking komt en die kenmerkend zijn voor het beroep, die aangepakt moeten worden binnen een specifieke organisatorische context, waarbij van de beroepsbeoefenaar een oplossing en een passende aanpak worden verwacht’. Elk kwalificatieniveau bestaat uit een aantal kernopgaven die ontwikkeld zijn door ontwikkelaars van het LSOP en vertegenwoordigers uit de politiekorpsen. Voorbeelden van kernopgaven zijn:

- kernopgave 2.1.2: Optreden bij milieuovertreding;
- kernopgave 3.3.2: Arrestantenzorg;
- kernopgave 4.1.1: Toezicht houden in publiek domein;
- kernopgave 5.4.3: Aanpak van verkeersveiligheid;
- kernopgave 6.2.4: Voorbereiding strategische aanpak bij rampen en calamiteiten;
- kernopgave 43.3.02: Organiseren en uitvoeren van opsporingsconfrontatie;
- kernopgave 50.0.03: Verkeersveiligheidsprojecten;
- kernopgave 66.6.06: Integraal managen van gemeenschappelijke veiligheidszorg.

Iedere kernopgave bevat een beschrijving van competenties die een student moet verwerven om de kernopgave uit te kunnen voeren. De competenties zijn door de ACOA ingedeeld in (Grotendorst et al., 2002):

- *Vakmatige en methodische competenties*. Deze vormen de kern van het beroepsprofiel en zijn gericht op het vermogen om op adequate wijze producten en diensten te leveren.
- *Bestuurlijk-organisatorische en strategische competenties*. Deze zijn gericht op het plannen en regelen van het eigen werk in de context van de organisatie.
- *Sociaal-communicatieve en normatief-culturele competenties*. Deze hebben betrekking op het functioneren in een werkomgeving, op samenwerking en op coördinatie van arbeid.
- *Leer- en vormgevingscompetenties*. Deze verwijzen naar het vermogen bij te dragen aan de eigen ontwikkeling en die van arbeidsorganisatie en beroep.

4.1.1 Toezicht houden in publiek domein

Resultaatgebied	Leefbaarheid
Voorbeelden	Winkelcentrum Markt
Werkpatroon (activiteiten)	1. Toezicht houden (ook op risicoplekken) 2. Observeren 3. Contacten leggen en onderhouden met publiek en collega's 4. Zelfstandig corrigerend optreden w.o. repressief optreden 5. Risico voor mogelijke escalatie inschatten, dit communiceren en corrigerend (laten) optreden 6. De weg wijzen 7. Eerste opvang slachtoffers
Voorbeeld-situatie	Er is een weekmarkt in het winkelcentrum. De allround politiemedewerker wordt door de assistent politiemedewerker geïnformeerd dat hij zojuist heeft geconstateerd dat er zakken worden gerold en op aangegeven signalement iemand heeft aangehouden. De assistent politiemedewerker kan dit niet zelfstandig afhandelen. De allround politiemedewerker zorgt voor onderlinge afstemming tijdens het door hem ingestelde opsporingsonderzoek.
Bekwaamheden	
Vakmatig-methodisch	Is in staat: – zich zodanig op te stellen dat hij actief aanwezig en aanspreekbaar is voor publiek: 'kennen en gekend worden' – waar te nemen en overzicht te houden met inachtneming van eigen veiligheid en die van anderen – afwijkend c.q. probleemgedrag te signaleren – ter plaatse de juiste informatie te kunnen verzamelen (vragen stellen en zelf waarnemen) – voorbeeldfunctie te vervullen – te handelen conform relevante wet- en regelgeving (proportioneel/discretionair) – gesprekstechnieken te gebruiken – door houding en voorkomen gezag uit te stralen – de uitgangspunten van gemeenschappelijke veiligheidszorg in zijn handelen te tonen – bij het ontstaan van incidenten: conflicten te hanteren, tussen verschillende partijen, ook bij grotere mensenmassa; indien nodig creatieve oplossingen te vinden – Integrale Beroepsvaardigheid Training (IBT) toe te passen (aanhoudingsprocedures en geweldtoepassing) – door houding en voorkomen gezag uit te stralen
Bestuurlijk-organisatorisch	Is in staat: – verantwoording af te leggen over zijn eigen handelen en in geval van gezamenlijk optreden over de met collega's gemaakte samenwerkingsafspraken – in te spelen op incidenten die zich voordoen en hierbij zijn taak uit te voeren in samenhang met de taken die anderen uitvoeren (zowel politie als anderen) – waar te nemen of, en in welk opzicht, andere politiemedewerkers begeleiding nodig hebben en is in staat deze begeleiding te bieden
Sociaal-communicatief	Is in staat: – bij wijzigingen in de situatie flexibel om te gaan met afspraken en dit terug te koppelen – met respect om te gaan met mensen van verschillende geaardheid en culturele achtergronden – zich dienstbaar op te stellen – duidelijkheid en standvastigheid te tonen
Leer- en vormgeving	Is in staat: – zijn eigen handelen te evalueren en op basis van verkregen inzichten dit bij te stellen – zijn beroepskennis en vaardigheden bij te houden.

Figuur 1. Een deel van kernopgave 4.1.1 'Toezicht houden in publiek domein'

Kwaliteitsgarantie	
Eerlijkheid	– opdracht geeft student de kans gedrag te tonen – examiner gerespecteerd door studenten
Openheid	– duidelijke informatievoorziening vooraf – doel assessment duidelijk voor studenten – status van uitkomst is duidelijk – te beoordelen competenties vooraf bekend bij studenten
Complexiteit	– combinatie van verschillende opdrachten – verschillende soorten competenties komen aan de orde – onderdelen geven ruimte voor diversiteit in gedragsalternatieven – te beoordelen competenties komen in meerdere opdrachten aan de orde
Authenticiteit	– opdrachten zijn gerelateerd aan relevante beroepssituaties/contexten
Directheid	– opdracht roept beroepsgerelateerd gedrag op – oordeel examinatoren is direct gerelateerd aan geobserveerd gedrag
Transparantie	– duidelijkheid over wat op welk moment beoordeeld wordt – eenduidig(e) beoordelingskader/criteria – criteria in beschrijvende schalen – duidelijke relatie tussen competenties en vertoond gedrag
Integratie leer-proces	– voor studenten herkenbare relatie met leerproces – kwalitatieve feedback als essentieel onderdeel
Zorgvuldigheid	– goede organisatie van de uitvoering – selectie en training van examinatoren – examinatoren niet tegelijkertijd begeleiders van studenten
Kwaliteitscontrole	
Doelmatigheid	– evaluatie studenten en examinatoren
Betrouwbaarheid	– beoordelaarbetrouwbaarheid – generaliseerbaarheid
Validiteit	– generaliseerbaarheid – begripsvaliditeit

Figuur 2. Criteria voor kwaliteitsgarantie en kwaliteitscontrole (naar Bruijns, 2002)

In figuur 1 geven we een deel van een kernopgave weer van de opleiding allround politiemedewerker: 4.1.1 ‘Toezicht houden in publiek domein’. In dit artikel nemen we deze kernopgave als voorbeeld.

Fase 3: Proeve van bekwaamheid

Na de vaststelling van de beroepsprofielen en de kernopgaven volgt de derde fase van de ontwerpmethodiek: het ontwikkelen van de proeve van bekwaamheid.

De proeve van bekwaamheid is het examen dat studenten in de gelegenheid stelt om de door hen verworven competenties in hun onderlinge samenhang te demonstreren, waarbij elke proeve gekoppeld is aan een kernopgave (Biemans et al., 2000). In deze fase wordt vastgesteld op welke wijze (door middel van welk type toetsituatie en examenopdrachten) de student demonstreert dat hij de competenties uit de betreffende kernopgave heeft verworven en op basis van welke criteria de beoordeling plaatsvindt. De toetskundigen

van het CEP hanteren een bewerking van de methodiek van Doorten en Moerkerke (1999, zie ook Bie-mans et al., 2000) die een aantal stappen kent, met als resultaat een proeve van bekwaamheid met examen-opdrachten, examenonderdelen en beoordelingscrite-ria.

KWALITEITSGARANTIE EN -CONTROLE

Hoe zorgt het CEP ervoor dat de proeve van bekwaamheid een vorm van deugdelijk assessment is? Met ande-re woorden: met welke kwaliteitscriteria houden toets-kundigen en examinatoren rekening tijdens de ontwikkeling en uitvoering van de proeven van bekwaamheid?

In de klassieke testtheorie beschouwt men validiteit en betrouwbaarheid als kwaliteitsindicatoren van de beoordeling. De psychometrische indicatoren zijn niet automatisch toepasbaar op nieuwe toetsvormen, zoals de proeven van bekwaamheid, omdat competentiebe-oordeling niet past binnen de klassieke testtheoreti-sche basis (Diederick, Van de Watering en Muijtens, 2002). In plaats van uitsluitend achteraf meten en waarderen van een activiteit of product, komt de nadruk meer te liggen op het vooraf vaststellen en zo veel mogelijk garanderen van de inhoudelijke en pro-cesmatige kwaliteit (Straetmans en Sanders, 2001; Bruijns, 2002). Figuur 2 (op de vorige pagina) geeft criteria aan die van invloed zijn op de kwaliteitsgaran-tie en kwaliteitscontrole.

Hieronder beschrijven we welke maatregelen het CEP neemt tijdens de ontwikkelings- en uitvoeringsfase ter bevordering van de kwaliteit. Hierbij maken we gebruik van de kwaliteitscriteria van Bruijns (2002).

ONTWIKKELING

Ontwikkelgroep

Voor de ontwikkeling van de proeven van bekwaamheid zijn ontwikkelgroepen samengesteld, bestaande uit vertegenwoordigers van de diverse disciplines. Naast de materiedeskundigen en trajectbegeleiders³ uit de korpsen zijn dit onderwijskundigen, docenten van het LSOP en Regionale Opleidings Centra en toetskundigen van het Centraal Examenbureau Politie (CEP).

De aanwezigheid van competenties is waarneembaar tijdens het uitvoeren van beroepsspecifieke activiteiten in een zo realistisch mogelijke, kritische beroepssitu-atie, waarbij de kandidaat beïnvloed wordt door een groot aantal factoren. De politieke vertegenwoordigers

in de ontwikkelgroepen geven aan welke kenmerken-de en relevante beroepssituaties gewenst gedrag (methoden, procedures, vakspecifieke aanpak en hou-ding) oproepen (authenticiteit en directheid).

Vervolgens bepaalt de ontwikkelgroep wanneer er sprake is van adequaat handelen van een student. De ontwikkelgroep stelt de beoordelingscriteria vast in ter-men van waarneembaar gedrag of productkenmerken. De multidisciplinaire ontwikkelgroep heeft op deze wijze positieve invloed op de validiteit van een proeve van bekwaamheid: de opdrachten zijn representatief en relevant voor de beroepspraktijk. Het is de taak van de toetskundige om het ontwikkelproces op zodanige wijze te sturen dat de proeve ook aan alle overige kwa-liteitsaspecten voldoet.

In competentiegericht politieonderwijs zijn de kern-opgave en de proeve van bekwaamheid leidend voor de inrichting van het leertraject (LSOP, 2001). Pas nadat de proeve van bekwaamheid bekend is, komt fase 4 uit de ontwerpmethodiek aan de orde: de ont-wikkeling van leeromgevingen waarin de student mogelijkheden krijgt de competenties te verwerven. Dezelfde ontwikkelgroep is vervolgens ook betrokken bij deze vierde fase. Hierdoor is de relatie tussen het beoordelingsproces en het leerproces gewaarborgd. Een gewenst neveneffect van de multidisciplinaire ont-wikkelgroepen is dat in de verschillende disciplines draagkracht gecreëerd wordt voor deze onderwijsver-nieuwing.

Methodenmix

Het CEP gebruikt de volgende methoden voor com-petentiebeoordeling (onder andere uit Straetmans, 1998; Korte, 2002):

- *Arbeidsproef*. Een student voert onder normale werk-omstandigheden een realistische opdracht uit (voor-beeld: toezicht houden op een markt).
- *Simulatie*. Een student voert onder gesimuleerde werkomstandigheden een realistische opdracht uit (voorbeeld: verkeersongeval). Hierbij kan het gaan om simulatie met acteurs, met apparatuur/instru-mentarium of computer.
- *Authentieke opdracht*. Dit betreft complexe en van de beroepspraktijk afgeleide probleemsituaties, waarin de student activiteiten ontplooit, resulterend in pro-ducten (voorbeeld: een student schrijft een plan van aanpak aan de hand van complexe casuïstiek rond-om een calamiteit of ramp).
- *Pen-en-papiertoets*. Hierbij gaat het om beoordelen van kennis en inzicht door middel van open of gesloten vragen (voorbeeld: meerkeuzetoets voor het meten van wetkennis).

Alle bovengenoemde methoden hebben zowel sterke als zwakke toetstechnische kanten. De arbeidsproef stemt het beste overeen met de uitgangspunten van competentiebeoordeling, maar is minder betrouwbaar (Korte, 2002). Dit heeft te maken met het feit dat deze competentiebeoordeling contextgevoelig is. Er is geen garantie dat de student in een andere situatie op dezelfde wijze wordt beoordeeld (generaliseerbaarheid). De pen-en-papiertoets kent een hoge betrouwbaarheid, maar een lage representativiteit (Korte, 2002).

Straetmans en Sanders (2001) geven aan dat er geen beste methode voor competentiebeoordeling is, maar dat een combinatie van toetsvormen, een methodenmix, uitkomst biedt. De methodenmix zorgt voor compensatie van zwakke punten van een bepaalde methode door de sterke punten van andere methoden. Daarnaast is meer gevarieerde informatie beschikbaar voor de beoordeling van bekwaamheden van een student.

Het is onmogelijk alle competenties uit een kernopgave te verwerken in één enkele examenopdracht. Om die reden bestaat een proeve van bekwaamheid uit meerdere examenopdrachten. Iedere examenopdracht kan daarbij bestaan uit meerdere examenonderdelen, die onderling met elkaar samenhangen. Bij deze examenopdrachten en examenonderdelen zorgt het CEP voor variatie in toetsvormen, voor compensatie van nadelen met andere toetsvormen en voor betrouwbare informatieverzameling over de bekwaamheid van een student. Dit past binnen het kwaliteitscriterium betrouwbaarheid, generaliseerbaarheid en complexiteit.

In figuur 3 geven we een voorbeeld van twee examenopdrachten uit één proeve van bekwaamheid van kernopgave 4.1.1 'Toezicht houden in publiek domein'. Examenopdracht 1.4 houdt in dat de student, samen met een andere student, een uur surveilleert in een winkelcentrum, op een station voor openbaar vervoer of in een woonwijk. Een examiner van het CEP (de assessor) volgt de studenten op korte afstand en legt de scores vast op een beoordelingsformulier.

PROEVE VAN BEKWAAMHEID 4.1.1 'TOEZICHT HOUDEN IN PUBLIEK DOMEIN'

Examenopdracht 1.0: Aanhouding- en zelfverdedigingstechnieken (simulatie)

In deze opdracht toon je tijdens een simulatie aan dat je de aanhouding- en zelfverdedigingstechnieken beheerst (zie art. 1 onder e van de Regeling Toetsing Geweldsbeheersing Politie, niveau initiële opleiding politiesurveillant). Je voert deze simulatie uit tijdens de instituutperiode. Je wordt beoordeeld door een (gecertificeerde) docent op het instituut, aan de hand

van een vastgesteld beoordelingsformulier.

Om te mogen deelnemen aan examenopdracht 1.4 uit deze proeve moet je examenopdracht 1.0 met een voldoende afsluiten.

Examenopdracht 1.4: Toezicht houden (arbeidsproef)

Deze opdracht houdt in dat je samen met een andere kandidaat gaat surveilleren in een publiek domein. Je bepaalt zelf in welk gebied je gaat surveilleren. Het gebied moet voldoen aan de gestelde eisen, die hieronder zijn genoemd. Als je deze opdracht uitvoert in een ander gebied dan waar je geoefend hebt, overleg je met je coach waar je deze opdracht het beste kunt uitvoeren.

Tijdens het toezicht houden wordt van je verwacht dat je de vier onderstaande activiteiten uitvoert. Hiermee neem je zelf het initiatief. Gedurende de gehele examentijd word je geobserveerd door de examiner van het CEP.

1. Je spreekt een persoon aan en je geeft hem een preventieadvies voor een kleine criminaliteit.
2. Je houdt een repressief (waarschuwend) gesprek naar aanleiding van een gepleegde overtreding of gedraging.
3. Je wijst de weg aan een persoon die hierom vraagt.
4. Je voert een gesprek met een slachtoffer (eerste opvang) van een eenvoudig misdrijf. Hierbij dient de nadruk te liggen op de wijze waarop je het slachtoffer opvangt en doorverwijst.

Figuur 3. Voorbeeld van simulatie en arbeidsproef uit proeve 4.1.1 'Toezicht houden in publiek domein'

Naast validiteit en betrouwbaarheid hangt de keuze voor de toetsvorm af van veiligheid van personen en situaties, ethische en organisatorische afwegingen. Het is niet wenselijk om een arbeidsproef af te nemen bij de aanhouding van een vuurwapengevaarlijke verdachte (veiligheid). Evenmin is het wenselijk om een arbeidsproef af te nemen bij een zedendelict of een daadwerkelijk verkeersongeval (ethische afweging). Een complex verkeersongeval is niet van tevoren te plannen. De simulatie verdient in dit geval de voorkeur boven een arbeidsproef (organisatorische afweging). Soms is gekozen voor een arbeidsproef met gesimuleerde onderdelen. In ons voorbeeld (figuur 3) kan het voorkomen dat er geen slachtoffer is van een eenvoudig misdrijf (bijvoorbeeld portemonnee gestolen). In dat geval wordt een acteur ingeschakeld om dit slachtoffergesprek met de kandidaat te voeren. De examenopdracht dient zodanig ingericht te zijn dat de student de kans krijgt om het beroepsgerelateerde gedrag te tonen. Dit zijn maatregelen gericht op eerlijkheid en directheid.

Beoordelingscriteria

Het oordeel van examinatoren over competenties van kandidaten is een complexe aangelegenheid en is gerelateerd aan direct observeerde of indirecte gedragsindicatoren of producten, waarvoor bepaalde eisen zijn gesteld. Om op een objectieve en betrouwbare wijze competenties te beoordelen, moet voor alle betrokkenen (van tevoren) bekend zijn waar het CEP op beoordeelt. Om deze reden zijn, in samenspraak met materiedeskundigen en praktijkvertegenwoordigers, voor elke examenopdracht beoordelingscriteria vastgesteld. Beoordelingscriteria worden niet te gedetailleerd en directief geformuleerd, om ruimteverkleining voor gedragsalternatieven te voorkomen. Het CEP geeft een toelichting op de beoordelingscriteria om een eenduidige interpretatie van de criteria te realiseren (zie figuur 4). Dit zijn maatregelen gericht op transparantie.

UITVOERING

Sturend voor leerproces

Zoals eerder aangegeven zijn de kernopgave en de proeve van bekwaamheid leidend voor de ontwikkeling van de leeromgevingen en sturend voor het leerproces van studenten. Een student is zelf verantwoordelijk voor zijn leerproces en maakt dus individuele keuzes in deze leeromgevingen. Daarom moet de proeve van bekwaamheid bij de start van de betreffende kernopgave bij de student bekend zijn. Dit is in tegenstelling tot het traditionele onderwijs, waarin het examen meestal pas aan het eind van een module of periode bekend is. De student ontvangt de beschrijving van zowel competenties uit de kernopgave als de beschrijving van de examenopdrachten. Hierin staat expliciet aangegeven wat, onder welke omstandigheden en in welke situatie van de student wordt verwacht bij de uitvoering van de examenopdrachten en op basis van welke criteria de beoordeling plaatsvindt (zie figuur 3). Tevens staat vermeld welke hulpmiddelen ter beschikking staan bij de uitvoering van de examenopdrachten (transparantie en openheid).

Erik: 'Ik deed examenopdracht 1.4 "toezicht publiek domein" op het Centraal Station in Den Haag. Tijdens de voorbereiding hierop merkte ik steeds weer dat de proeve sturend is voor mijn leertraject. Ik hield steeds in de gaten wat van mij verwacht werd aan het einde van de kernopgave. Waar misschien meer rekening mee gehouden zou kunnen worden, is het examenmoment; ik deed om 10 uur examen, ja toen was het niet zo druk dat ik kon laten zien wat ik in huis heb.'

De proeve is als het ware een resultaatafspraken tussen de student, de docent op het instituut en de coach in de praktijk. Om de competenties behorende bij kernopgave 4.1.1 'Toezicht houden in publiek domein' te verwerven, zet de kandidaat zich gedurende 10 weken in op het opleidingsinstituut om vervolgens 10 weken in de praktijk aan de slag te gaan. In deze periode bieden respectievelijk de docent en de coach expertise aan om de student te ondersteunen bij de competentieverwerving. In het publieke domein bepaalt de student zelf een gebied waar hij een surveillance kan uitvoeren, gelet op de speerpunten die in dat gebied van belang zijn voor een toezichthouder of politieagent. Tijdens deze voorbereiding maakt de student gebruik van leeropdrachten vanuit het onderwijs en examenopdrachten (inclusief beoordelingscriteria) vanuit de proeve. De proeve van bekwaamheid helpt bij het geven van gerichte feedback door de docent en praktijkcoach. Deze maatregelen zijn gericht op integratie met het leerproces, openheid en transparantie.

Examinatoren

Er is in het vernieuwde politieonderwijs gekozen voor een onafhankelijke beoordeling van competenties door het CEP (LSOP, 2001). Het CEP werkt op dit moment met negen vaste examinatoren en ongeveer veertig freelance examinatoren. Deze examinatoren zijn zelf executief politieambtenaar, dat wil zeggen dat zij jarenlange ervaring hebben als politieagent en inhoudsdeskundig zijn op het gebied van de basispolitiezorg of een specialisme. Idealiter blijven ze werkzaam in de praktijk. De examinatoren verdienen hierdoor het respect van studenten als beoordelaar (eerlijkheid). Naast de examinatoren van het CEP kunnen ook praktijkcoaches, trajectbegeleiders en lijnfunctionarissen als beoordelaar optreden. Zij mogen echter niet betrokken zijn bij de beoordeling van studenten die zij zelf hebben begeleid. Onafhankelijke beoordeling past bij het kwaliteitscriterium 'zorgvuldigheid'.

Bart Rijkenberg (fulltime examiner van het CEP): 'Was ik destijds zelf maar zo geëxamineerd; de examens zijn erg praktijkgericht; het publiek heeft vaak niet of nauwelijks in de gaten dat het een examen betreft.'

Martin Fontijne (fulltime examiner van het CEP): 'Omdat je als examiner zelf ook gewerkt hebt in de basispolitiezorg, ben je goed op de hoogte van de gang van zaken op straat. Het is voor ons erg belangrijk contact te houden met de ontwikkelaars van de proeven, zodat we ook de ideeën achter de proeve kennen. We doen ook voorstellen aan de toetskundigen om de proeven te optimaliseren.'

Vanessa: 'De examinatoren vroegen op de examenlocatie aan collega's wat de plaatselijke gebruiken waren. Je merkt dat de examiner zelf het politievak beheerst.'

Beoordelingsinstructie en -formulier

De examinatoren van het CEP zijn getraind in competentiebeoordeling. Daarna volgt een training per proeve van bekwaamheid om de uitvoering te stroomlijnen en om de beoordelaarbetrouwbaarheid te verhogen. Hiervoor is per proeve van bekwaamheid door toetskundigen een *uitvoeringsdocument* geschreven, waarin de procedure van de afname van de examenopdrachten is weergegeven. Tevens staat in dit document beschreven op basis van welke criteria de beoordeling plaatsvindt, inclusief een toelichting op deze beoordelingscriteria (zie figuur 4).

3. OPTREDEN:

Algemene toelichting:

- de gesprekken moeten zo gevoerd worden dat het beoogde doel bereikt wordt, waarbij de klanttevredenheid relatief is (zal bij de weg wijzen groter zijn dan bij het waarschuwingsgesprek)
- voor alle gesprekken geldt dat vriendelijkheid en doortastendheid er zo veel mogelijk in door dienen te klinken
- bij het waarschuwingsgesprek dient de methode 'Van der Steen' te worden gehanteerd (direct meedelen dat het een waarschuwing is)

Checklist beoordelingscriteria:

3.1 geven van preventieadvies t.a.v. 'kleine criminaliteit':

- toepassen gesprekstechnieken
- gespreksdoel aangeven, logisch advies geven (gerelateerd aan risicofactoren), beargumenteerd, duidelijk, niet moraliserend
- gevaren zien, potentiële slachtoffers opmerken, in huid kruipen van verdachte, relateren aan voorkomende modus operandi

3.2 voeren van waarschuwingsgesprek:

- b.v. fietsen in voetgangersgebied, parkeerovertreding, APV-feit
- toepassen gesprekstechnieken
- gespreksdoel aangeven, waarschuwing is relevant, beargumenteerd, duidelijk, niet moraliserend
- juist toepassen wet- en regelgeving, controleren of maatregel wordt uitgevoerd
- kennen sociale kaart
- toepassen methode 'Van der Steen'

Het uitvoeringsdocument bevat de vertrouwelijke informatie over de examenopdrachten. Bij een simulatie betreft dit bijvoorbeeld de beschrijving voor de rollenspelers, of de antwoordsleutel voor een authentieke opdracht. Het uitvoeringsdocument is om die reden uitsluitend bedoeld voor de examiner en/of rollenspelers. Het uitvoeringsdocument past tezamen met de training bij de kwaliteitscriteria 'betrouwbaarheid' en 'zorgvuldigheid'.

Het CEP ontwikkelt per examenopdracht een beoordelingsformulier waarop de criteria staan weergegeven (zie figuur 5). Het gebruik van dit beoordelingsformulier heeft als voordeel dat de beoordeling door examinatoren vergemakkelijkt wordt (zorgvuldigheid). De beoordelingsformulieren zijn eenduidig, waardoor alle examinatoren de kandidaten op dezelfde punten beoordelen. Dit heeft een positieve invloed op de betrouwbaarheid van de beoordeling.

Het beoordelingsformulier is een handzaam en werkbaar instrument voor de examiner. Een ingewikkeld en uitvoerig formulier kan een goede observatie door de examiner belemmeren, omdat het invullen van het formulier te veel tijd en concentratie vergt. Daarom is ervoor gekozen om het aantal criteria te beperken én de criteria op volgorde van handelen te plaatsen op het formulier.

Scoring en cesuur

Bij het ontwerpen van de beoordelingsformulieren kiest het CEP voor de tweepuntsschaal, omdat dit het meest betrouwbaar is. Omdat een tweepuntsschaal weinig feedback voor de kandidaat bevat, is op de beoordelingsformulieren ruimte gecreëerd waar de examiner tijdens of na afloop van het examen feedback geeft. De examiner is verplicht om een onvoldoende score altijd te voorzien van beargumentering. Na de examenopdracht volgt een gesprek waarin de uitslag van de proeve bekend wordt gemaakt en welke onderdelen wel of niet voldoende zijn gescoord.

Bepaalde competenties uit een kernopgave zijn essentieel voor een goede uitvoering van de examenopdracht. Studenten moeten een voldoende scoren op deze cruciale beoordelingspunten die direct gerelateerd zijn aan deze essentiële competenties. Een beoordelingscriterium is cruciaal wanneer de onvoldoende uitvoering daarvan leidt tot het ernstig benadelen van betrokken partijen of tot vormfouten. De cruciale beoordelingscriteria zijn in een aparte kolom op het beoordelingsformulier (zie figuur 5) weergegeven. Ieder cruciaal beoordelingscriterium moet voldoende scoren. Van de overige, niet-cruciale punten geldt een cesuur van 70 procent. Van dit percentage kan worden afgeweken, indien dit van tevoren is vastgesteld en op voorhand bekend is bij de student.

Figuur 4. Toelichting enkele beoordelingscriteria behorende bij examenopdracht 1.4

 LSPD • Centraal Samenwerkend Politie Delict, Delict- en Criminologie		Beoordelingsformulier Proeve 4.1.1 Toezicht houden publiek domein Examenopdracht 1.4: Toezicht houden	
Naam	J. Jansen	Evenementcode	
Studentnummer		Examencode	
Cursuscode		Examinator	
Locatie van afname	Amsterdam	Tijdstip	:
Datum	01 - 01 -	Eindoordeel	☐ Onvoldoende ☒ Voldoende

Checklist	Cruciaal	O	V	VC	Verantwoording
1 Voorbereiding					
1.1 informatie verzamelen	☒	☐	☐	☐	
1.2 surveillanceplan opstellen	☐	☐	☐	☐	
1.3 overige voorbereidingen treffen	☐	☐	☐	☐	
1.4 zelfstandig werken	☐	☐	☐	☐	
2 Toezicht houden					
2.1 zichtbaar aanwezig, aanspreekbaar zijn	☐	☐	☐	☐	
2.2 observeren, signaleren	☐	☐	☐	☐	
2.3 inschatten, communiceren	☐	☐	☐	☐	
2.4 beslissen en dit motiveren	☒	☐	☐	☐	
2.5 actief uitnodigend zijn (initieel)	☐	☐	☐	☐	
3 Optreden					
3.1 geven preventieadvies	☐	☐	☐	☐	
3.2 voeren waarschuwingsgesprek	☐	☐	☐	☐	
3.3 de weg wijzen	☐	☐	☐	☐	
3.4 voeren opvanggesprek slachtoffer	☐	☐	☐	☐	
3.5 vriendelijk zijn	☐	☐	☐	☐	
3.6 doortastend zijn	☐	☐	☐	☐	
4 Verbindingsmiddelen					
4.1 gebruik verbindingsmiddelen	☐	☐	☐	☐	
5 Beroepshouding					
5.1 professionele houding	☒	☐	☐	☐	

Alle cruciale punten dienen voldoende te worden gescoord. Van de niet cruciale punten moeten er minimaal 10 voldoende zijn. Is dit niet het geval dan is het examenresultaat onvoldoende.

Algemene opmerkingen / bijzondere omstandigheden

Opmerking 1: kandidaat verwijst naar verkeerde verkeersregel (fietsen op voetpad) en controleert niet of de maatregel wordt uitgevoerd.

Handtekening kandidaat voor deelname aan het examen

Handtekening examiner

Voorbeeld beoordelingsformulier

Figuur 5. Fictief beoordelingsformulier van examenopdracht 1.4 'Toezicht houden'

Bevoegd en inzetbaar

Een voldoende resultaat op de proeve van bekwaamheid (dus aangetoonde bekwaamheid) leidt tot certificering en formele bevoegdheid en inzetbaarheid in de korpsen. Een student die de proeve 4.1.1 'Toezicht in het publiek domein' heeft afgesloten, wordt vervolgens in de praktijk daadwerkelijk ingezet voor toezichtactiviteiten in het publiek domein.

Kwaliteitscommissie en examenraad

Naast de eerder genoemde maatregelen, hanteert het CEP instrumenten voor interne en externe kwaliteitszorg. Wat betreft interne kwaliteitszorg kent het CEP een kwaliteitscommissie, bestaande uit toetskundigen van het CEP. Deze beoordeelt de proeven van bekwaamheid op de toetskundige en onderwijskundige kwaliteit. De verantwoordelijke toetskundige ontvangt van de kwaliteitscommissie aanbevelingen om de kwaliteit en eenduidigheid van de producten te verhogen. De externe kwaliteitszorg is de taak van de Examenraad van het CEP, bestaande uit externe deskundigen op het gebied van onderwijs en politie. De Examenraad organiseert audits waarin de proeven worden beoordeeld op basis van kwaliteitscriteria. Naar aanleiding van deze audit stuurt de Examenraad een rapport van bevindingen over de kwaliteit van de proeven naar het College van Bestuur van het LSOP Politie Onderwijs- en Kenniscentrum.

LEERERVARINGEN

De ontwikkeling van de proeven van bekwaamheid is in volle gang. De eerste proeven van bekwaamheid zijn afgenomen. In deze paragraaf staan we stil bij een aantal leerervaringen van het CEP tijdens de ontwikkeling en uitvoering van de proeven van bekwaamheid.

Financiën

De wijze van examineren is ingrijpend veranderd. De uitvoer van de proeve van bekwaamheid is op individuele basis. Dit heeft consequenties voor de organisatie van examinering met de daarmee samenhangende financiering. Tevens kost de ontwikkeling van de proeven van bekwaamheid (bijvoorbeeld de uitwerking en uitvoering van een simulatie) veel tijd en geld. Om deze redenen moet het CEP binnen de gestelde kaders steeds keuzes maken.

Competentiebeschrijving

De beschrijving van de competenties behorende bij de kernopgaven zijn te algemeen geformuleerd voor de ontwikkeling van examenopdrachten. Een competentie van kernopgave 4.1.1 is bijvoorbeeld: 'Een student is in

staat om te handelen conform relevante wet- en regelgeving'. Deze competentie komt ook in alle overige kernopgaven naar voren. De competenties staan namelijk los van de context, terwijl deze context belangrijk is voor het ontwikkelen van goede examenopdrachten. Om meer inzicht en eenduidigheid te verkrijgen over de competenties hebben de ontwikkelgroepen de kernopgaven geanalyseerd. De analyse maakt duidelijk wat als basis nodig is voor de te hanteren methoden en procedures. Wat betreft het eerder genoemde voorbeeld is duidelijk in kaart gebracht welke wetteksten en regelgeving relevant zijn voor de betreffende kernopgave.

Integratie en overlap

Zoals hiervoor is aangegeven, komen een aantal competenties bij meerdere kernopgaven terug. Dit sluit aan bij de competentiegerichte gedachte (generaliseerbaarheid: competent in verschillende situaties). Voor de ontwikkeling van proeven is het van belang om vast te stellen op welk moment in het leertraject de student deze (overlappende) competenties op het juiste niveau verworven heeft, zodat deze kunnen worden beoordeeld. Als bij één proeve blijkt dat een student voldoende bekwaam is, moet dezelfde competentie dan bij de volgende proeven steeds opnieuw terugkomen?

Context van competentiebeoordeling

Omdat de beoordeling van een student op individuele basis plaatsvindt, speelt de context een belangrijke rol. Nederland kent verschillende korpsen die gekoppeld zijn aan arrondissementen⁴. Zij moeten voldoen aan landelijk vastgestelde richtlijnen. Ieder korps heeft echter een bepaalde speelruimte binnen deze richtlijnen (bijvoorbeeld het formulier voor een proces-verbaal). Daarnaast is de problematiek in regio's niet zonder meer vergelijkbaar. Deze regionale verschillen hebben invloed op de beoordelingscontext en hierdoor kunnen verschillen ontstaan in de zwaarte van de examenopdracht voor verschillende studenten. Zo stelt toezicht houden op een markt in een dorp andere eisen dan toezicht houden op een markt in de drukke binnenstad.

Gedragsindicatoren

Sommige competenties zijn moeilijk te meten met behulp van direct observeerbare gedragsindicatoren, waardoor het oordeel gebaseerd is op indirect waarneembare gedragsindicatoren en/of producten. De methodenmix van Straetmans en Sanders kan in dit geval zorgen voor een evenwicht tussen direct en indirect waarneembare gedragsindicatoren. Het CEP past deze methodenmix steeds meer toe in de proeven van bekwaamheid.

Kwaliteitscontrole

De eerder genoemde maatregelen zijn met name gericht op de kwaliteitsgarantie, maar minder gericht op kwaliteitscontrole. Het CEP richt zich nu naast de kwaliteitsgarantie ook op maatregelen met betrekking tot kwaliteitscontrole. Op dit moment ontwikkelt het CEP een instrument voor gestructureerde studentevaluatie. Samen met de evaluatiegegevens van de examinator worden de proeven van bekwaamheid aangepast voor het volgende cohort studenten (doelmatigheid). Het CEP overweegt maatregelen op het gebied van psychometrische kwaliteit van de beoordelingsformulieren (betrouwbaarheid). Tevens overweegt het CEP bij de werving en selectie van examinatoren gebruik te maken van een proeve van bekwaamheid voor examinatoren (betrouwbaarheid).

PROEVE VAN BEKWAAMHEID ALS DEUGDELIJK ASSESSMENT

De verwachtingen van de 'nieuwe' studenten, die het LSOP volgens de competentiegerichte onderwijsbenadering opleidt, zijn hooggespannen. De eerste reacties van studenten, docenten, korpsen, examinatoren en toetskundigen van het CEP met betrekking tot competentiebeoordeling door middel van de proeve van bekwaamheid zijn positief. Vertegenwoordigers uit de korpsen herkennen nu al verschillen tussen oude en nieuwe studenten. De proeve van bekwaamheid 4.1.1 'Toezicht houden in publiek domein' kent een slaggingspercentage van 100 procent (gegevens bekend tot en met oktober 2002). Op de vraag naar meer blauw op straat geven we een passend antwoord, namelijk: competent blauw op straat!

Wij hopen dat dit artikel u inzicht geeft in de proeve van bekwaamheid als deugdelijk assessment. We sluiten dit artikel af met suggesties die u kunt gebruiken voor de ontwikkeling van uw eigen deugdelijke instrumenten voor competentiebeoordeling:

- *Ontwikkeling*. Ontwikkel het assessment samen met een multidisciplinaire groep docenten, praktijkvertegenwoordigers en toetskundigen. Hierdoor ontstaan beroepsrelevante, toetstechnisch adequate opdrachten en beoordelingscriteria. Zorg voor een breed draagvlak voor de competentiebeoordeling.
- *Methodenmix*. Zorg voor variatie in opdrachten. Hierdoor kun je voldoende betrouwbare informatie verzamelen voor competentiebeoordeling. Gebruik een mix van arbeidsproeven, simulaties, authentieke opdrachten en pen-en-papiertoetsen.
- *Randvoorwaarden*. Een student moet bij het assessment de gelegenheid krijgen om de bekwaamheden te bewijzen. Houd er rekening mee dat niet alles mogelijk is. Zorg voorafgaand aan de ontwikkeling voor duidelijkheid over de randvoorwaarden zoals veiligheid, organiseerbaarheid en financiering.
- *Proeve bekend*. Maak het assessment aan het begin bekend, zodat een student weet wat hij kan verwachten. Dit betreft zowel de opdrachten als de beoordelingscriteria. Hierdoor werkt het assessment sturend op het leer- en/of ontwikkelingsproces.
- *Assessoren*. Laat competentiebeoordeling plaatsvinden door ervarings- en/of materiedeskundigen. Laat de examinatoren de bekwaamheid in competentiebeoordeling bewijzen.
- *Beoordelingsinstructie*. Zorg voor een duidelijke beoordelingsinstructie in de vorm van training voor de afname van het assessment en een toelichting op de beoordelingscriteria.
- *Beoordelingsformulier*. Ontwikkel een hanteerbaar beoordelingsformulier met adequate beoordelingscriteria.
- *Kwaliteitscontrole*. Zorg naast kwaliteitsgarantie ook voor kwaliteitscontrole achteraf.

MEER INFORMATIE

Meer informatie over de proeven van bekwaamheid voor het politieonderwijs is te verkrijgen bij het Centraal Examen bureau Politie van het LSOP, Postbus 186, 7300 AD Apeldoorn, tel. 055-5392200, fax 055-5392315.



Literatuur

- Adviescommissie Onderwijs-Arbeidsmarkt (ACOA), *Een wending naar kerncompetenties: De betekenis van kerncompetenties voor de versterking van de kwalificatiestructuur secundair beroepsonderwijs*. Den Bosch: ACOA, 1999.
- Biemans, H., F. de Jong, C. Hoogendoorn en A. Grotendorst, *Handleiding bij de proeven- en onderwijsontwikkeling PO2002*. Apeldoorn: LSOP Politie Onderwijs- en Kenniscentrum, 2002.
- Bruijns, V., 'Assessment centers als onderdeel van een competentiegerichte leeromgeving'. In: F. Dochy, L. Heylen en H. van de Mosselaer (red.), *Assessment in onderwijs. Nieuwe toetsvormen en examinering in studentgericht onderwijs en competentiegericht onderwijs* (pp. 161-180). Utrecht: Uitgeverij Lemma bv, 2002.
- Dierick, S., G. van de Watering en A. Muijtjens, 'De actuele kwaliteit van assessment: ontwikkelingen in de edumetrie'. In: F. Dochy, L. Heylen en H. van de Mosselaer (red.), *Assessment in onderwijs. Nieuwe toetsvormen en examinering in studentgericht onderwijs en competentiegericht onderwijs* (pp. 91-122). Utrecht: Uitgeverij Lemma bv, 2002.
- Doorten, M. en G. Moerkerke, *Performance Assessment: meten en beoordelen van complexe vaardigheden*. Open Universiteit Nederland, 1999.
- Grotendorst, A., M. Jellema, I. Stam, M. van der Vegt en C. Zandbergen, *Leren in veiligheid. Het nieuwe politieonderwijs in maatschappelijk perspectief*. Apeldoorn: LSOP Politie Onderwijs- en Kenniscentrum, 2002.
- Kessels, J.W.M., *Succesvol ontwerpen: Curriculumconsistentie in opleidingen*. Deventer: Kluwer Bedrijfswetenschappen, 1996.

- Korte, M., *Proeven van bekwaamheid: De ontwikkeling van competentiegerichtte examens voor het vernieuwde politieonderwijs*. Apeldoorn: LSOP Politie Onderwijs- en Kenniscentrum, 2002.
- LSOP Politie Onderwijs- en Kenniscentrum, *Functioneel ontwerp van Politieonderwijs 2002: De inrichting van het vernieuwde politieonderwijs* (oktober 2001). Apeldoorn: LSOP Politie Onderwijs- en Kenniscentrum, 2001.
- Straetmans, G.J.J.M., *Toetsing van competenties*. Arnhem: CITO, 1998.
- Straetmans, G.J.J.M. en P.F. Sanders, *Beoordelen van competenties van docenten*. Den Haag: Programmamanagement EPS/HBO-raad, 2001.

Noten

1. De overige fasen uit de ontwerpmethodiek worden niet in dit artikel behandeld.
2. Het politieberoep kent geen beroepsprofiel dat overeenstemt met kwalificatieniveau 1 (Korte, 2002).
3. Trajectbegeleiders scheppen voorwaarden voor effectief leren en bewaken de voortgang van studenten tijdens het werkend leren.
4. Dit zijn rechtbankgebieden waar politiekorpsen de processen-verbaal naar toezenden om vervolging van strafbare feiten in gang te zetten.

Met dank aan V. Koolmees, E. van Herk, M. Fontijne, B. Rijkenberg, A. Grotendorst, L. Dercksen en M. Korte.



KANTTEKENINGEN BIJ EEN HYPE

HOGESCHOLEN HEBBEN WEINIG OOG VOOR DE CONSEQUENTIES VAN INTAKE-ASSESSMENTS

ANJA GELDERMANS

5

Steeds meer hogescholen starten met intake-assessments. Door aankomende studenten te beoordelen op (elders verworven) competenties is het mogelijk onderwijstrajecten op maat aan te bieden. Het leveren van individuele trajecten blijkt in de praktijk echter maar mondjesmaat van de grond te komen. 'Het onderwijs zit muurvast op groepsgroottes en standaardgroepen. Studenten kunnen hooguit een aantal vrijstellingen krijgen en daar houdt het mee op.' Ook het onderzoek naar de validering van de uitkomsten van intake-assessments blijft op grote schaal achterwege. 'We weten nog steeds niet of de resultaten iets zeggen over iemands latere functioneren.'

Een prima beoordeling kreeg Carla Eichholtz voor het assessment dat zij dit voorjaar aflegde voor het zij-instroomstraject van de lerarenopleiding. Eichholtz, die jarenlang werkzaam was als medewerker bestrijding onderwijsachterstanden bij de Bibliotheek Amsterdam, meldde zich enige tijd geleden bij de Stichting Lerarentekort Amsterdam (nu de Onderwijs bv). Haar vorige functie stond door personele onderbezetting voortdurend onder druk en een baan als docent leek haar een goede overstap. De assessmentprocedure vond ze heel nuttig. 'Een assessment is honderd keer beter dan een sollicitatiegesprek,' zegt Eichholtz. 'Je wordt gedwongen goed over je keuze na te denken. Vooral het schrijven van een portfolio dwingt je tot reflectie.'

Eichholtz was verbaasd over de inhoud van het assessment. 'Het eerste onderdeel bevatte een taal- en reken-toets. De taaltoets bestond slechts uit zestig spellingsvragen. Ik dacht nog: de Nederlandse taal heeft toch wel meer om het lijf?' Vervolgens moest ze een proefles draaien op een basisschool. Eichholtz: 'Voor een kleuterklas maakte ik een les over lieveheersbeestjes. Ik had me goed voorbereid en omdat ik al ervaring heb met het maken van lesmateriaal verliep de les prima.' Na afloop kreeg ze een eindgesprek met de intern begeleider van de basisschool en een docent van de lerarenopleiding. De beoordelaars waren enthousiast over haar prestaties en deelden Eichholtz mee dat zij de tweejarige opleiding wellicht in één jaar zou kunnen afronden. Met deze uitslag stapte Eichholtz naar de Educatieve Faculteit Amsterdam. Maar een maattraject kon de lerarenopleiding haar niet bieden, deelde de onderwijscoördinator mee. Eichholtz zou zelf moeten bewijzen dat ze in staat is de opleiding in één jaar te doen. 'Maar hoe kan ik dat aantonen?' Vervolgens probeerde ze het bij de Hogeschool Ipabo in Amsterdam. 'Ik kreeg een vrouw aan de lijn die begon te lachen. Ook daar bleek het niet mogelijk een individueel traject te volgen. Ik vind dat heel jammer. Ze beweren maatwerk te kunnen leveren, maar kunnen dat absoluut niet waarmaken.' Ondanks haar teleurstelling is ze dit studiejaar toch begonnen met de opleiding. 'Ik doe het uiteindelijk voor mijn eigen ontplooiing.'

Steeds meer hogescholen en opleidingen starten met intake-assessments. Via het intake-assessment kan worden vastgesteld over welke (elders verworven) competenties een student beschikt. Vervolgens kunnen opleidingen aan de hand van de uitkomsten de student een aangepast opleidingstraject aanbieden. Pabo's en lerarenopleidingen zijn, onder druk van het lerarentekort, al enkele jaren geleden met intake-assessments begonnen. Maar ook (deeltijd)opleidingen in andere sectoren bieden steeds vaker assessments aan bij de poort. Hogescholen kunnen op deze manier nieuwe doelgroepen binnenhalen en beter inspelen op de behoeften van de arbeidsmarkt. Voor studenten wordt het onderwijs aantrekkelijker: hun leertraject wordt korter en bevat precies de leerstof waar ze behoefte aan hebben.

Om een maattraject te kunnen aanbieden dat perfect is afgestemd op de kwalificaties van de student, is het belangrijk dat een intake-assessment betrouwbaar is. Het opstellen van een assessment dat betrouwbaar en valide is, blijkt echter geen sinecure. 'De ontwikkeling vraagt veel meer tijd dan mensen er vooraf voor uit willen trekken,' merkt Harm Tillema, hoofddocent onderwijskunde aan de Universiteit Leiden. 'Maar als je als hogeschool een assessmentbureau inricht, heren der leentjebuurt speelt en daarmee een quick assessment in elkaar draait, dan gaat het fout. Zo'n initiatief vertaalt zich direct door in allerlei vervelende uitkomsten.' Ook Erik Roelofs, onderwijskundige en methodoloog bij het Cito, stelt dat het proces om competenties in kaart te brengen niet eenvoudig is. 'Het moeilijke is dat competenties erg contextspecifiek zijn. Iemand kan bijvoorbeeld een expert zijn in het lesgeven aan kleuters, maar totaal niet uit de verf komen bij andere groepen. De uitkomsten van competentiemetingen kun je dus niet zomaar generaliseren. Je moet daarom voldoende verschillende situaties en instrumenten meenemen bij de metingen.'

Nog moeilijker wordt het als toekomstige studenten worden beoordeeld op 'elders verworven competenties' (EVC), zoals bij de zij-instromerstrajecten voor de lerarenopleidingen. Roelofs: 'De relatie tussen algemene competenties die je eerder hebt opgedaan en het presteren in een nieuwe situatie wordt in onderzoek niet duidelijk aangetoond. Het is dus moeilijk om te voorspellen of iemand uit een houtzagerij of een herintredende moeder goed zal presteren in het latere beroep. Bij deze assessments wil je dus eigenlijk iets anders meten: heeft iemand talent en zou hij of zij zich in de toekomst kunnen ontwikkelen tot een goede docent? Eigenlijk is het heel simpel: als je veel vergelijkbare werkervaring hebt opgedaan, heb je de meeste kans om te slagen.'

Een andere reden waarom er weinig hoge voorspellende waarden bij assessments worden gevonden, ligt volgens Roelofs aan het feit dat assessments altijd kunstmatige situaties zijn. De omstandigheden waaronder mensen hun taken moeten uitvoeren blijven hierbij buiten beeld. Roelofs: 'Het is sterk de vraag hoe representatief deze situaties zijn voor de latere beroepspraktijk. Eigenlijk zou je in assessments iets moeten meenemen van de tijdsdruk, de complexiteit van de situatie en de randvoorwaarden die het werk bemoeilijken. Je kunt bijvoorbeeld goed zijn in organiseren, maar ben je dat ook met dertig kinderen om je heen?'

MOOIE WOORDEN

Roelofs ergert zich aan alle mooie woorden over een belangrijk onderdeel van assessments, het portfolio. 'Het portfolio wordt een instrument genoemd, maar het is niets anders dan een verzameling van bewijsmaterialen. Er worden geen eenduidige eisen gesteld aan de inhoud en het is vaak onduidelijk of je het ene portfolio kunt vergelijken met het andere. Digitale portfolio's zijn bovendien vaak veel te omvangrijk. Sommige beslaan wel honderden pagina's.' Roelofs voorkeur gaat uit naar compacte dossiers van niet meer dan vijftien A-viertjes. 'Ellenlange verhalen zijn niet interessant. Je moet dus vooraf een screening maken van het materiaal.'

Het assessment zelf mag vooral niet te veel tijd in beslag nemen. Roelofs: 'De meting moet betrouwbaar zijn, maar ook praktisch uitvoerbaar. Je moet dus zoeken naar een zo eenvoudig mogelijk instrumentarium dat zo goed mogelijke uitkomsten geeft. Observeren is bijvoorbeeld interessant, maar je moet wel tientallen observaties doen, wil je een fatsoenlijk beeld van iemand krijgen.'

Als je honderd procent betrouwbaarheid nastreeft, moet je misschien wel twee weken toetsen, zegt ook Erica Ter Wee, onderzoekster bij Stoas Onderwijs. 'Maar negentig procent is toch ook goed?' Ter Wee ontwikkelde enkele jaren geleden – in samenwerking met Harm Tillema – in opdracht van het ministerie van onderwijs een procedure om de geschiktheid van zij-instromers voor het beroep van leraar vast te stellen. Ter Wee somt een lijst van eisen op waar een goed assessment nog meer aan moet voldoen: transparant, acceptabel, duidelijk, valide en niet te duur. 'Vaak bijten deze eisen elkaar.'

Tillema en Roelofs wijzen daarnaast beiden op de belangrijke rol van de assessor. Roelofs: 'Je moet ervoor zorgen dat een beoordelaar niet geblindeerd is door de eigen voorkeuren en daardoor zijn eigen ideeën projecteert. Je kunt dus het beste met meerde-

re assessoren werken, zodat zij elkaar kunnen corrigeren. Ook moet je als assessor geen afstreeplijstjes hanteren. Op die manier maak je geen gebruik van je meerwaarde als beoordelaar.' Tillema stelt dat gedrag alleen goed kan worden beoordeeld door mensen die werkplekdeskundigheid hebben. 'De assessor moet dus een vakman zijn in het werk én een vakman in beoordelen. Docenten die willen werken als assessor doen er daarom goed aan zich te laten trainen.'

Zit het gezien al deze eisen wel goed met de kwaliteit van de intake-assessments bij hogescholen? 'Vraag je dat aan mij?', zegt Tillema. 'Ik heb al jaren geleden gezegd dat we moeten kijken naar de kwaliteit. In de

De assessor moet een vakman zijn in het werk én in beoordelen

euforie waarin iedereen op competenties wilde beoordelen, is men gaan experimenteren met nieuwe vormen van assessments. Dat is prima in de beginfase, maar op een gegeven moment komt de vraag of deze metingen wel voldoen.' Ter Wee maakt zich inmiddels zorgen over de uitvoering van de assessments die door het Stoa's zijn ontwikkeld. 'De lerarenopleidingen hebben de procedure naar eigen inzicht verder ingevuld, waarbij zij onderdelen hebben veranderd of vervangen door andere.' De kwaliteit is er daarna niet altijd beter op geworden, meent zij. 'Op dit moment is de onderwijsinspectie bezig de praktische uitvoering bij de verschillende instellingen nader te bekijken.'

Inmiddels heeft Tillema op initiatief van de lerarenopleidingen een audit ontwikkeld, waarmee hogescholen zelf de interne controle kunnen verzorgen op de kwaliteit van assessments. In de audit zijn 460 kwaliteitskenmerken verzameld. Beoordelaars kunnen afvinken of bepaalde elementen wel of niet aanwezig zijn. Aan de hand van de uitkomsten kunnen de opleidingen verbeteringen doorvoeren. Tillema bezocht in de periode tussen maart 2001 en januari 2002 zeven lerarenopleidingen. Uit de audits bleek dat de assessments goed worden uitgevoerd. 'Dat vond ik heel prettig, maar ik stuitte wel op een ander probleem. Het blijkt dat opleidingen het assessment zien als een aparte problematiek, dat ergens door een bureauje wordt uitgevoerd. Ze nemen assessments af, maar doen verder niets met de resultaten. Zo raakt het geïsoleerd. Nu is het vaak een afzonderlijke, losstaande beoordeling die niet verder verwijst. Daardoor ervaren docenten het assessment als een last die niets oplevert.'

Ter Wee vindt dat hogescholen niet goed nadenken over de consequenties van competentiegericht beoor-

delen. 'Als je uitgaat van individuele kwaliteiten van mensen, dan moet je ook individuele leertrajecten verzorgen. Dat is onlosmakelijk met elkaar verbonden.' Ze constateert dat hogescholen niet in staat zijn handvatten te vinden voor de individualisering van het onderwijs. 'Het onderwijs zit muurvast op groeps-groottes en standaardgroepen. Het lijkt erop alsof opleidingen denken dat maatwerk niet mogelijk is en het daarom ook maar niet proberen. Zij-instromers kunnen wel vrijstellingen krijgen, maar daar houdt het mee op. Het door ons ontwikkelde assessment is echter zo ontworpen dat als iemand aan alle competenties voldoet, hij of zij direct een diploma kan krijgen. In de praktijk doen de scholen echter aan koppelverkoop. Ze koppelen altijd een opleiding aan het assessment. Dat is niet eerlijk tegenover de kandidaten.'

DOOD IN DE POT

Een van de problemen bij het aanbieden van flexibele leertrajecten is dat kandidaten tijdens een assessment worden beoordeeld op competenties, maar dat het curriculum niet competentiegericht is. Voor docenten blijkt het heel moeilijk om het idee van competenties te vertalen naar de onderwijsinhoud, constateert Tillema. 'Docenten proberen altijd competenties te vertalen naar vakken. Dat is natuurlijk de dood in de pot, daar zijn competenties niet voor bedoeld.'

Ter Wee: 'Als hogescholen geen competentiegericht onderwijs aanbieden, dan levert de aansluiting met het intake-assessment een gekunstelde situatie op. Competentiegericht onderwijs biedt meer mogelijkheden voor flexibele leerroutes.'

Hogescholen kunnen dit probleem niet blijven negeren, denkt Tillema. 'Op een gegeven moment keert de wal het schip. De instroom bij hogescholen wordt steeds heterogener. Je kunt dan niet meer volstaan met reguliere voltijd- of deeltijdopleidingen. Je zult steeds gevarieerdere onderwijstrajecten moeten aanbieden. En je kunt studenten geen assessments afnemen, zonder dat het voor hen zichtbaar nut heeft. Op een gegeven moment krijg je het van de gebruikers terug: waar dient dit eigenlijk voor? Wat kan ik ermee?'

En dat terwijl intake-assessments de opleiding en de student juist veel voordelen kunnen opleveren. Ter Wee: 'Als je als hogeschool in staat bent om flexibele leertrajecten aan te bieden, kun je veel profijt halen uit het intake-assessment. Het leertraject wordt aantrekkelijker voor studenten, omdat ze precies dat krijgen waar ze behoefte aan hebben. Daardoor zijn ze beter gemotiveerd.'

Tillema wijst op een andere belangrijke kwestie die nog niet is opgepakt: de validering van de uitkomsten. Meten we inderdaad wat we willen meten? En hebben

de resultaten van het intake-assesment voorspellende waarde? Tillema: 'We weten nog niet of de uitkomsten van intake-assessments iets zeggen over iemands latere functioneren. Onderzoek hiernaar mist nog volstrekt, maar is hard nodig. Ik heb dit herhaaldelijk aangekaart, zowel bij het ministerie als bij de hogescholen, maar niemand wil er geld of tijd voor vrijmaken. Toch moet dat onderzoek worden opgepakt voordat we ons hele onderwijs op assessments baseren.'





ADVISERING RONDON SUCCES:

VERANTWOORDELIJKHEDEN, VALKUILEN EN MOGELIJKHEDEN MET
BETREKKING TOT DE TOEPASSING VAN ASSESSMENTS

ANTHON VAN DER HORST

6

Wie zich met het beoordelen van mensen bezighoudt in de context van hun functioneren in arbeidsrelaties neemt een grote verantwoordelijkheid op zich. Ik wil aan de hand van mijn ervaringen toelichten wat deze verantwoordelijkheid behelst. Daarbij wil ik me richten op eenieder die mensen beoordeelt in arbeidsrelaties, ook op degenen die geen gebruikmaken van de assessment-centermethode. Bij hen gaat het erom aan te geven welke risico's zij nemen door er juist geen gebruik van te maken. Degenen die er wel gebruik van maken, wil ik laten zien op welke wijze zij er in mijn ogen mee om moeten gaan, wanneer ze extra goed moeten opletten. Verder wil ik betogen dat het gebruikmaken van de systematische beoordeling van mensen de gebruiker extra mogelijkheden geeft om zijn eigen organisatie in bredere zin te begrijpen, maar ook om specifieke doelgroepen scherper in kaart te krijgen om ze vervolgens beter te kunnen managen. Het onderstaande komt voor een belangrijk deel voort uit mijn ervaring, de lezer moet het artikel dan ook zien als opiniërend, de eigen gedachten aanscherpend, en niet als het laatste ware woord.

HISTORIE

Selecteren en beoordelen is van alle tijden. Een vroeg voorbeeld hiervan vinden we in het Oude Testament, waar Gideon een keurkorps samenstelt door driehon-

derd mannen uit een beek te laten drinken. Degenen die knielden en direct met hun mond uit de beek dronken, werden geselecteerd voor de opdracht. Het ging in dit geval om het beoordelen van het vertrouwen in hun meester. Als die zei dat je het water kon drinken, hoefde je niet meer te twijfelen en kon je je onvoorwaardelijk overgeven aan de kwaliteit van het water. Wie eerst het water zelf checkte, toonde te weinig vertrouwen en was in de ogen van de meester ongeschikt voor de strijd waar onvoorwaardelijk vertrouwen noodzakelijk was.

Zo zijn er nog meer voorbeelden voorhanden. De eerste serieuze psychologische test is door Alfred Binet ontworpen. De Franse overheid wilde namelijk een methode om meer en minder begaafde kinderen van elkaar te kunnen onderscheiden om zo de aanwezige onderwijsmiddelen optimaal te kunnen inzetten. De eerste echt grootschalige toepassing van de psychologische test vond plaats in de Eerste Wereldoorlog, naar het schijnt voor een belangrijk deel vanuit efficiencyoverwegingen. Het was makkelijker om schriftelijk vragen te stellen en te verwerken dan iedereen persoonlijk te interviewen. Het assessment center is sterk opgekomen in de Tweede Wereldoorlog. In de VS werden spionnen opgeleid die in een vierdaags assessment center onder andere praktijksituaties moesten doen. Zo moesten zij bijvoorbeeld aantonen dat ze zich konden redden in de situatie waarin ze

gesnapt werden door de Duitsers op verboden terrein (Drenth, 1985).

BEGRIPSBEPALING

Voor de definiëring van wat een assessment is, neem ik een ruime marge. Uiteindelijk is een assessment letterlijk genomen een meting en zo wil ik het ook graag zien. Een meting bij mensen van competenties die nodig zijn voor het succesvol functioneren in een functie. Als het goed is, vindt deze meting plaats met behulp van wetenschappelijk verantwoorde meetmethoden. Dat betekent vanuit mijn referentiekader een meting die bestaat uit een mix van oefeningen (simulaties) die gedrag oproepen (vaak assessmentoefeningen genoemd), psychologische testen (persoonlijkheidsvragenlijsten en intelligentietesten), maar ook het criteriumgerichte interview.

Omdat assessment letterlijk ‘meting’ betekent, zouden we alle overige vormen van meten van menselijke competenties een vorm van assessment kunnen noemen, maar daar kies ik niet voor. Vooral vanuit pragmatische overwegingen. Binnen mijn praktijk is een assessment meestal, zo niet altijd, een combinatie van hiervoor genoemde meetmethoden.

WERKWIJZE

Hoe gaat een goed assessment in z’n werk? Ik ga in op de analyse van competenties, aandachtspunten voor de rapportage en het belang van het zorgvuldig omgaan met de verantwoordelijkheden door alle betrokkenen.

Competentieprofilering

Een goed assessment wordt voorafgegaan door een zorgvuldige analyse van de omstandigheden waarin iemand komt te werken. Een veelgebruikte methode is die van competentieprofilering, waarbij met behulp van kritische incidenten, een analyse van de bedrijfsstrategie, van de omstandigheden waarin iemand moet functioneren, een set van gedrag wordt geformuleerd waaraan iemand moet voldoen om goed te functioneren. Goed functioneren is daarbij gedefinieerd door de opdrachtgever, en dat is vaak de baas van de functionaris of een HRM-functionaris.

Het assessment: rapportage

Het vervolg is dat de deelnemer wordt uitgenodigd om zich vervolgens meestal een dag onder te dompelen in allerhande rollenspellen, testen en interviews. Binnen ongeveer een week wordt het rapport met de deelnemer besproken. Deze heeft het recht op eerste inzage volgens de beroepscode van de beroepsvereniging van het NIP (Nederlands Instituut voor Psychologen) en

kan indien dat wordt gewenst ook besluiten dat het rapport niet aan de opdrachtgever wordt doorgestuurd. Als het rapport bij de opdrachtgever is beland, kan er nog nader overleg plaatsvinden, waarbij de psycholoog er wel zeer goed op moet letten dat hij hetzelfde vertelt als in het rapport. Nooit kan een psycholoog in zo’n gesprek een andere inkleuring geven als tijdens zijn rapportage. Het streven is dan ook dat het rapport een zo duidelijk antwoord geeft op de vragen van de opdrachtgever dat nader overleg niet meer nodig is.

Verantwoordelijkheden bij de terugkoppeling

De beoordelaar heeft bij de terugkoppeling een grote verantwoordelijkheid. Immers, de kandidaat geeft zich behoorlijk bloot in de onderzoekssituatie, heeft daar vaak van tevoren allerlei fantasieën over, bijvoorbeeld als de beoordelaar een bedrijfspsycholoog is, dat deze ‘dwars door de kandidaat heen kijkt’, en het is dan ook zeer belangrijk dat de privacy gewaarborgd is en de gevoeligheden van de deelnemer gerespecteerd en erkend worden.

Wie niet aan een beroepscode gebonden is, zal een hele duidelijke gedragscode moeten hanteren om diezelfde privacy toch te waarborgen. Fouten worden al snel gemaakt. Zo maken wij vaak mee dat opdrachtgevers de dag na het assessment al bellen en vragen wat de uitslag is. Op dat moment moet je aangeven dat de kandidaat als eerste recht heeft op de uitslag. Opdrachtgevers vinden dat niet altijd prettig! Een collega van mij zwichtte een keer en gaf in zo’n telefoongesprek aan dat de opdrachtgever zich geen zorgen hoefde te maken over de uitslag, daarmee implicerend dat het een positief advies zou zijn. De opdrachtgever bleek meteen daarna de kandidaat te hebben gebeld om hem te feliciteren met de positieve uitslag van de test. Deze kandidaat heeft uiteindelijk geen klacht bij het NIP ingediend, maar had dat zeker wel kunnen doen en zijn klacht zou zijn gehonoreerd. Uit dit voorbeeld blijkt ook dat de opdrachtgever een grote verantwoordelijkheid heeft. De sollicitant meet de zorgvuldigheid van een werkgever vaak af aan de wijze waarop deze omgaat met dit soort procedures. Een onzorgvuldige afhandeling van de werkgever van het assessment leidt soms ook tot terugtrekking van de kandidaat voor de functie.

Tot slot: wat betreft de werkwijze zal er in de praktijk ook regelmatig sprake zijn van evaluatie tussen opdrachtgever en beoordelaar, bedrijfspsycholoog. Dat kan met behulp van een informeel gesprek, maar ook met behulp van wetenschappelijk onderzoek. Evaluatieonderzoek kan tot doel hebben de succesfactoren nog beter in kaart te krijgen, maar ook om de selectie efficiënter te maken. In de paragraaf ‘Succesvoorspelling’ geef ik hier enkele voorbeelden van.

HEDENDAAGSE PRAKTIJK

In het bedrijfsleven is AT&T het eerste bedrijf dat jonge managers ging assen en in Nederland is Mars een van de eerste bedrijven die het assessment center introduceerde. Nu zullen er in Nederland jaarlijks vele tienduizenden mensen een assessment of een vorm van assessment ondergaan. Het assessment center is in Nederland dus sterk ingeburgerd. Eigenlijk is het eerder bijzonder als je nog nooit een vorm van assessment of psychologische test hebt meegemaakt, dan wanneer je het al wel hebt meegemaakt. In mijn praktijk komt het voor dat mensen al vier of vijf keer een assessment hebben ondergaan!

De historie laat zien dat zeker in Nederland het instrument vaak en breed wordt toegepast. Het is in eerste instantie voortgevloeid uit de behoefte om fouten te

- De potentiële waarde van mensen voor de onderneming zou vollediger gerealiseerd kunnen worden dan meestal het geval is.
- HRM maakt een wezenlijk onderdeel uit van het algehele strategische beleid van een onderneming en dient met andere beleidsterreinen te worden afgestemd.
- HRM is een lijnverantwoordelijkheid, zowel aan de top als voor midden- en lager management.
- De medewerker is niet alleen voorwerp van HRM, maar heeft een duidelijke verantwoordelijkheid voor de eigen ontwikkeling (Hoekstra en Van Sluis, 1999).

Wat deze uitgangspunten betekenen voor een organisatie is in mijn ogen het volgende. Wie HRM-instrumenten als assessment wil gaan toepassen, maakt een duidelijke keuze voor de benadering van zijn mensen, of dat nu medewerkers zijn of studenten. Het betekent dat het management zich veel meer verdiept in het wel en wee van de medewerker, zich faciliterend tot de ontwikkeling van de medewerkers opstelt, en goed begrijpt wat succes in de eigen organisatie veroorzaakt. Tevens weet het management dit succes te vertalen naar individuele functies. Competentiemanagement is daarbij het gevleugelde woord. De verantwoordelijkheid van de manager, maar ook van de adviseur die HRM serieus neemt, is deze samenhang tussen succes en competentie management te zien en te vertalen naar de praktijk. Het werken met assessments is één zo'n vertaling of toepassing, maar zeker niet de enige. Een goede beoordelingssystematiek hoort daarbij, coaching, opleiding, enzovoort.

TOEGEVOEGDE WAARDE VAN ASSESSMENTS

Niets is zo lastig als het beoordelen van mensen en niets geeft managers zo veel zorgen als disfunctionerende medewerkers. En voor de meeste bedrijven in Nederland bepalen de mensen niet alleen de inkomsten maar ook de kosten. Een foute aanname kost een bedrijf al snel veel geld. Toch past niet iedereen het assessment toe. Ook zijn er nog andere bezwaren die worden geuit tegen assessment. Hieronder schets ik de vier meest genoemde, waarbij mijn opvatting is dat de verantwoordelijke manager en de verantwoordelijke deelnemer deze bezwaren eigenlijk niet zouden moeten hebben.

Assessments zijn duur!

De verantwoordelijkheid voor de juiste selectie van personeel houdt menig manager druk bezig. Stelt u zich voor: een vakgroepmanager van een school, waar vijf afdelingshoofden onder vallen en daaronder weer honderd leerkrachten. Deze manager heeft als

De kunst van het voorspellen van fouten biedt ook de mogelijkheid om succes te voorspellen

voorkomen, maar vlot werd ontdekt dat de kunst van het voorspellen van deze fouten ook de mogelijkheid biedt om succes te voorspellen en ontwikkelingsbehoeften vast te stellen. Zo worden psychologische testen en assessments al snel ook gebruikt in de context van loopbaanvragen, sterkte-zwakteanalyses, de beroepskeuzetest, enzovoort.

De vormen waarin dat gebeurt variëren sterk. Er zijn in Nederland nog steeds grafologen actief op hetzelfde gebied, er zijn mensen die zogenoemde kleuren-testen gebruiken. Veel van deze testen, die ik graag exotische assessments noem, worden ook vandaag de dag nog toegepast. Wetenschappelijk gezien zijn deze methoden nooit goed te onderbouwen, maar wellicht zijn de normen die de wetenschap juist niet hanteert de normen die in deze exotische praktijken gelden. Zelf kom ik uit de traditie van het wetenschappelijk verantwoorde voorspellen, maar wel met een warme persoonlijke interesse voor al datgene wat werkt in de praktijk van het voorspellen.

ASSESSMENT: ONDERDEEL VAN EEN GOED HRM-BELEID

Het assessment is een van de mogelijkheden die een manager of een afdeling P&O heeft om het personeel beter te managen. Kort wil ik hier ingaan op wat de kern van HRM is, omdat het veel zegt over de context van mijn ervaringen:

- Menselijke kwaliteiten vormen een strategische succesfactor in elke onderneming.

opdracht om de leiding van deze vakgroep te verzorgen, nieuwe lesmethoden te introduceren, alsmede het imago van de vakgroep in de regionale gemeenschap verder positief uit te bouwen. De school heeft begroot dat er een groei van 10 procent mogelijk moet zijn als de doelen slagen. Er is een budget van € 10 miljoen. De kosten van de afdeling bestaan voor 75 procent uit loonkosten. De manager verdient € 65.000. De lezer voelt al aankomen wat hier het afbreukrisico is, als het met deze functionaris mis gaat in deze context. Want stel dat deze manager door een gebrek aan visie, overtuigingskracht en coachende vaardigheden zijn team niet weet te motiveren tot veranderen, hij door onhandig manoeuvreren weerstanden creëert, waardoor afdelingsmanagers ziek worden, leerkrachten extra onder druk komen te staan, waardoor het verloop enkele procenten stijgt, wellicht zelfs met als eindresultaat dat de manager zelf gedesillusioneerd vertrekt met grote twijfel aan zijn eigen kunnen ... Om het helemaal dramatisch te maken, kan het zelfs zo zijn dat hij in een persoonlijke crisis geraakt, met alle privé-gevolgen van dien. Ook kan een dergelijke gang van zaken nog een heel ander effect hebben, want wie was eigenlijk verantwoordelijk voor de aannahme van deze manager? Hoe zit het met de positie van de directie, wat vindt de Raad van Toezicht ervan als dit soort dingen gebeuren? Niet alleen kan de schade enkele tonnen, zo niet miljoenen bedragen, het imago van de school heeft niet de gewenste verbetering doorgeemaakt, de directie kan hierdoor de hete adem van een ontevreden bestuur voelen, et cetera.

Met andere woorden: de zorgvuldige selectie met gebruikmaking van een assessment betekent in mijn ogen dat de opdrachtgever zich kostenbewust gedraagt.

Hoewel de impact minder groot zal zijn, geldt de kern van de redenering ook voor 'lagere functies'. De feitelijke kosten zijn relatief gering. In de markt zullen de tarieven variëren tussen de € 1.000 en € 3.500, afhankelijk van de complexiteit van de functie en de zwaarte van het onderzoek. Hoe groot het verlies is als het mis gaat, heb ik hiervoor aangegeven. Persoonlijk vind ik de kosten dan ook absoluut geen argument om niet aan assessments te doen.

De uitslag is beïnvloedbaar ...

In mijn praktijk hoor ik de scepticus nogal eens stellen dat de waarde van het assessment gering is, daar de kandidaat de uitslag kan beïnvloeden. Ik geef meteen toe dat de kandidaat de testresultaten kan beïnvloeden. Sterker: dat is ook de bedoeling, want het is zijn invloed die we juist interessant vinden! Maar bedoeld wordt natuurlijk dat de kandidaat de testresultaten zo kan beïnvloeden dat hij sociaal wenselijke antwoorden

kan geven met als doel de beoordelaar te misleiden, waardoor deze beoordelaar de kandidaat niet ziet zoals deze werkelijk is. De daarbij behorende aannahme is dat als de beoordelaar de misleiding ziet, hij de kandidaat zal afwijzen. Een in mijn ogen merkwaardige aannahme. De deelnemer kan er beter van uitgaan dat hij geschikt is, en dat dat er vanzelf uit zal komen door zichzelf te zijn.

Het beïnvloeden van een assessment in positieve zin is overigens lastig. Een deelnemer aan een assessment kan heel moeilijk inschatten wat nu precies het gewenste gedrag is voor een functie. De verschillende vragenlijsten zijn niet dermate transparant dat de consequentie van het positief of negatief beantwoorden van een vraag meteen duidelijk is. Een voorbeeld kan dit verduidelijken. Veel kandidaten denken dat het

Mijn advies aan deelnemers van een selectie-assessment is: zet je beste beentje voor, maar wel je eigen beste beentje!

tonen van onzekerheid in bijvoorbeeld een managementfunctie niet verstandig is. Daardoor laten ze zich zien van een stevige kant, tonen veel besluitvaardigheid, initiatief, tonen niet snel hun gevoel en beantwoorden alle vragen die te maken hebben met onzekerheid ontkennend. Het beeld dat dan zal ontstaan, is er een van een ijzeren Hein, een hard, weinig invoelelend, niet sensitieve persoon. Een beetje een botte boer. Welk bedrijf zoekt zo'n soort manager? Daar komt bij dat het voor een kandidaat ontzettend moeilijk is, zo niet ondoenlijk, om consequent een ander beeld van zichzelf neer te zetten. Laat staan de testresultaten zo te beïnvloeden dat ze net niet teveel van het slechte zijn en ook weer niet overdreven goed. De kandidaten die ik ontmoet geven negen van de tien keer aan zichzelf te zijn, en laten zien ook goed te begrijpen dat als ze een eerlijk beeld schetsen van zichzelf de uitkomsten ook in hun belang zijn. Namelijk: pas ik in de functie? De kandidaat die het wel probeert, zal snel merken dat de beoordelaar vindt dat er iets niet klopt, het roept een onbestemd gevoel op bij de beoordelaar, wat leidt tot twijfel. De gemiddelde manager redeneert 'bij twijfel niet binnenhalen', met als gevolg een afwijzing.

Ik vind dat kandidaten ook een verantwoordelijkheid hebben, zowel naar het bedrijf als naar zichzelf en hun naaste omgeving, namelijk de verantwoordelijkheid van eerlijkheid over zichzelf en hun kwaliteiten. Mijn advies aan deelnemers van een selectie-assessment is dan ook: zet je beste beentje voor, maar wel je eigen beste beentje!

We kennen hem al

Een derde, ogenschijnlijk zeer valide, argument om iemand niet te assessen, is het argument dat de betrokkene al bekend is bij het management. Het kan toch niet zo zijn dat we iemand die we al een x-aantal jaren kennen nog eens aan een externe beoordeling onderwerpen? Dat is toch wel een zwaktebod! Mijn antwoord luidt: Neen! Het is geen zwaktebod als de nieuwe functie daadwerkelijk verschilt van de vorige. Een binnen dienstmedewerker in een commerciële organisatie die vertegenwoordiger wil worden, die kan de manager nog zo goed kennen, maar hij zal niet weten hoe zo iemand het doet in deze geheel nieuwe rol. De beste verpleegkundige die teammanager wordt, is in die rol niet bekend bij zijn meerderen. Deze mensen zijn natuurlijk dezelfde, maar kunnen toch ineens heel ander gedrag vertonen dan verwacht. Wie kent de voorbeelden niet.

Het vraagt moed van het management om in een dergelijke situatie, daar waar er ook vaak een goede persoonlijke band is, toch de verantwoordelijkheid te nemen voor een objectiverend oordeel, met als doel een goede match tussen persoonlijke kwaliteiten en functie-eisen te bewerkstelligen.

Iemand kent zichzelf al

In het verlengde van dit laatste ligt ook het vierde argument dat bij een loopbaanonderzoek of een sterkte-zwakteanalyse gehoord wordt, namelijk het argument dat iemand zichzelf al kent en wat kan een dag onderzoek (zo lang duurt een onderzoek vaak) daar nog aan toevoegen? Ik gebruik dan vaak de metafoor van het kijken naar de maan. Wij allen kunnen bij heldere lucht naar de maan kijken. Wij zien dan allemaal min of meer hetzelfde, ervan uitgaande dat onze ogen goed zijn en anders onze bril. Wie met het blote oog naar de maan kijkt, ziet een vrij kleine witte bol met enkele vlekken en sommigen zullen er een gezicht in ontwaren. Wat gebeurt er echter als wij door een verrekijker naar de maan gaan kijken? Allereerst wordt de bol veel groter en zien we veel meer details, we kunnen verklaren waarom de vlekken vlekken zijn, we zien zelfs hoogteverschil. Kortom: diezelfde maan onthult ineens veel meer van zichzelf dan we in eerste instantie dachten.

Zo ook het assessment. Op de vraag 'Wat vindt u van uzelf?' of 'Wat zijn uw sterke en zwakte punten?' hebben de meeste mensen wel een antwoord. Ze gebruiken misschien vier of vijf woorden om positieve eigenschappen te beschrijven en twee woorden om negatieve eigenschappen te beschrijven. Dat zijn ze, vinden ze zelf.

Daarmee doen ze zichzelf schromelijk tekort, ze nemen niet de verantwoordelijkheid voor wie ze werkelijk zijn.

Een rapport dat ik maak over een kandidaat bevat gemiddeld zo'n honderd typering, nuances met daarnaast aanbevelingen hoe zich verder te ontwikkelen. Natuurlijk zitten daar ook typering bij die een deelnemer zelf ook noemt, maar de beschrijving die de beoordelaar kan geven, de nuance, het treffende detail, is voor de deelnemer vaak van grote waarde omdat deze net verklaart waarom hij of zij niet altijd goed overtuigt of waarom er omstandigheden zijn waaronder de druk te groot wordt. Een assessment is vergelijkbaar met het kijken door een grote telescoop, en dan niet naar de maan maar naar jezelf.

SUCCESVOORSPELLING

Het voorgaande kan de indruk wekken dat de adviseur die zich met assessments bezighoudt vooral bezig is bezwaren weg te nemen, en voortdurend anderen moet overtuigen waarom ze een assessment zouden moeten doen. Dat valt wel mee. In Nederland is het assessment vaak een geaccepteerd instrument. Het thema 'toegevoegde waarde van assessments' speelt altijd in de voorfase van een assessment en bij het beantwoorden van de vraag 'Waarom zou ik een assessment laten doen?' of 'Waarom zou ik een assessment ondergaan?'. Als deze vraag positief is beantwoord komen er nieuwe thema's aan bod, die ik hier graag bespreek: aandachtspunten voor een geslaagd assessment. De vijf thema's die ik nu aan bod stel, hebben te maken met het voorspellen van succes.

Succespsychologie

Als psycholoog realiseerde ik mij enkele jaren geleden ineens het volgende. De psychologie wordt in de volksmond vaak geassocieerd met problemen. Wie naar de psycholoog gaat die heeft een probleem, voelt zich vaak niet goed in zijn vel, toont zich in zekere zin zwak. Hoewel het ook in Nederland steeds gewoner wordt om naar een therapeut te gaan, heeft de gang naar de psycholoog een zweem van negativiteit. Het aardige van het vak dat de gemiddelde bedrijfspsycholoog (kan ook door andere disciplines worden beoefend) beoefent, is dat hij of zij zich bezighoudt met de psychologie van het succes oftewel met succespsychologie.

Iemand die aan mijn bureau zit voor een selectie-assessment heeft vaak al een goede baan, is door een opdrachtgever al uitgeselecteerd en vaak is de vraag: kan deze persoon in de functie succesvol worden? Voor betrokkenen is het vaak zo dat de vraag is: kan ik in deze zwaardere functie ook succesvol worden? De beoordelaar is dus vooral een succesbeoordelaar. Een zeer positieve context en eigenlijk een zeer vrolijk beroep. In de praktijk is deze benadering er een die alle partijen energie geeft en het element van een

strengere beoordeling, een soort examen eraf kan halen. Het thema is niet zozeer: is de kandidaat geschikt en als deze het niet is, is hij ongeschikt, maar: de kandidaat is al succesvol, wij vinden het belangrijk dat hij succesvol blijft en we gaan samen bepalen of deze nieuwe functie ook tot succes zal leiden.

Succes komt in vreemde gedaanten maar is altijd herleidbaar tot gedrag

Het beoordelingsvak beoefen ik nu ruim 13 jaar en ik heb in al die jaren ruim 3000 mensen geassesseerd. Mijn organisatie heeft in de loop van haar 55-jarig bestaan

Elk succes is herleidbaar tot gedrag en dit gedrag is vaak goed te meten

veel expertise opgebouwd. Een omgeving en een wereld waar de oplettende adviseur zeer veel kan leren. Onder andere heb ik er geleerd dat succes door opdrachtgevers vaak aan vreemde zaken toegeschreven wordt. Mijn stelling is echter dat elk succes herleidbaar is tot gedrag en dat dit gedrag vaak goed te meten is, alsmede dat het goed mogelijk is mensen daarin op te leiden. Het volgende voorbeeld illustreert deze stelling.

Voorbeeld commercieel succes

Ik heb alweer enige tijd geleden een onderzoekje gedaan naar het succes van vertegenwoordigers. Tijdens dit onderzoek sprak ik met een verkoopleider van een bedrijf dat producten aan de warme bakker verkocht. Hij stelde dat de succesvolle vertegenwoordiger bij voorkeur gereformeerd was, getrouwd was met een vrouw die niet werkte en een hypotheek had. U begrijpt daar had ik lang tegenaan kunnen assen. Maar na enig doordenken en doorvragen bleek dat het hem erom ging dat deze mensen de taal van de gereformeerde bakker spraken (het achterland was gereformeerd). Dat hij getrouwd was, betekende voor de verkoopleider dat hij een stabiel thuisfront had en zich volledig kon wijden aan zijn werk, en de hypotheek stond ervoor dat de vertegenwoordiger niet te snel rare dingen zou doen en betrouwbaar was, omdat hij anders door ontslag zelf in de financiële problemen zou komen. Het ging dus in termen van gedragscriteria om sociabiliteit, optreden, om prestatiemotivatie en discipline. En die competenties konden wij wel meten (Damen en Van der Horst, 1992)!

Succescriteria juist interpreteren

Niet alleen is het belangrijk om de juiste woorden te geven aan datgene wat tot succes leidt en vooral te bewerkstelligen dat meer mensen begrijpen wat iemand succesvol maakt, ook is het wezenlijk om succescriteria juist te interpreteren. Ik schets twee voorbeelden.

Voorbeeld 1: de retailer

Een klant van mij, een grote uiterst succesvolle retailer, selecteert medewerkers bij wie ondernemerschap hoog in het vaandel staat. Ik vertel u kort even iets over het bedrijf opdat duidelijk is waar het hier om gaat.

Wie bij dit bedrijf binnenkomt kwam tot voor kort in een bescheiden wachtruimte waar niet veel aan de muur hangt, maar wel een foto van de eerste winkel van de founding father van het bedrijf. Tot voor kort stond er ook een man die een denkbeeldige trekschuit voortrok. Het motto van deze onderneming had veel te maken met volharding, doorzettingsvermogen. Het was het beeld van zwoegen en ploeteren, recht door zee, niet al te wendbaar in combinatie met soberheid en een winkel. Voor de selectie van medewerkers gold het strikte uitgangspunt dat de kandidaat retailervaring moest hebben, daar werd niet van afgeweken. De kandidaat-ondernemer moest bij voorkeur zelf in een retailfamilie zijn grootgebracht. Het vervelende echter van deze eis was dat de retail nou niet het imago had waar iedereen zich door aangetrokken voelde. Daarbij was en is de formule van deze retailer dermate succesvol dat er veel nieuwe ondernemers nodig waren. Doordat de winkels steeds groter werden, was de complexiteit ook groot en moest er ook de opleidingseis gesteld worden van hbo, met name op cijfermatig gebied en organisatorisch inzicht. Het gevolg was dat er een tekort aan kandidaat-ondernemers ontstond.

Deze organisatie heeft onder andere twee dingen gedaan: zij hebben het imago van *Volharding in Groei* zeer serieus genomen en niet verward met starheid, maar goed ingezien dat om in deze groei te volharden er juist flexibiliteit nodig was. Zij hebben dat gedaan door met professor Kessels te gaan nadenken over dit probleem. De oplossing bleek eigenlijk het ei van Columbus. Simpel en daardoor door iedereen over het hoofd gezien. Nagegaan is wat er eigenlijk bedoeld werd met retailervaring, die ervaringskennis is geoperationaliseerd en omgezet in een opleiding, waardoor er ineens een veel grotere doelgroep aangeboord kon worden. De selectie-eis staat nog steeds, maar de aanpassing van de opleiding aan het ontbreken van de ervaringseis was een mooi staaltje van creatief en flexibel denken. Het letterlijke beeld van die winkel in de ontvangsthuis durfde men los te laten en men realiseerde zich dat het niet ging om de ervaring van huis

uit in een winkel, maar om de competenties die te maken hadden met alertheid, oordeelsvorming en ondernemerschap. De biografische eis ervaring met retail is anders geïnterpreteerd en vertaald naar gedrag en dat gedrag kon in een opleiding worden nagebootst. En het is dat gedrag, niet de achtergrond, dat tot succes leidt.

Voorbeeld 2: het gespecialiseerde advocatenkantoor

Een tweede voorbeeld dat ik geef is van een gerenomeerd gespecialiseerd advocatenkantoor. De praktijk van dit kantoor is gericht op fiscale advocatuur, alwaar het onder andere zeer belangrijk is dat de advocaat zeer precies kijkt naar de letter van de wet en de exacte procedures en daarbij behorende tijdslijnen. De voorzitter van de maatschap had onder andere het criterium 'omgang met details' opgegeven als een cruciaal succescriterium. Daarnaast zocht hij ook creatieve mensen die zich snel konden aanpassen aan veranderende omstandigheden. Die eigenschappen gaan in de praktijk eigenlijk niet goed samen, zo bleek in mijn onderzoeken. De kandidaten waren gericht op het detail, of ze waren aanpassingsbereid en creatief. Aanpassingsbereid en creatief gingen nog wel samen, maar de combinatie met detailgerichtheid werd niet aange troffen. Toch gaf deze opdrachtgever aan niet alleen zelf zo te zijn, maar ook enkele collegae om zich heen te hebben die ook aan deze eigenschappen voldeden. Na een intensieve discussie bleek dat succes vooral bepaald werd door een zeer sterke prestatiegerichtheid, een sterke drive om het goed te doen en dat daarvoor een grondige, kritische houding nodig was, zonder dat het noodzakelijk was dat iemand bovengemiddeld ordelijk, detailgericht is. Met andere woorden: een kritische, prestatiegerichte houding maakte het dat iemand geen fouten maakte, wat de opdrachtgever in eerste instantie vertaalde naar omgang met details, maar wat eigenlijk prestatiegerichtheid genoemd moest worden.

Succescriteria juist durven benoemen

Een belangrijk onderdeel van het benoemen van succescriteria is het verantwoordelijkheid nemen voor het *durven* benoemen van succescriteria. Werkelijk succes gaat in mijn ogen niet altijd gepaard met gedrag of persoonskenmerken waar onze ouders trots op zouden zijn geweest. Ik kan dat illustreren aan de hand van een groep bekende Nederlanders die ik heb mogen onderzoeken. Ik wil dit voorbeeld ook gebruiken om te laten zien wat de consequenties kunnen zijn als we deze succescriteria willen vertalen naar een opleidingssituatie.

Voorbeeld 3: de bekende Nederlanders

Onlangs ben ik gevraagd om een aantal bekende Nederlanders te testen. Eerst voor BNN, later voor

Yorin. In dit voorbeeld wil ik laten zien dat deze succesvolle mensen niet meteen het succesprofiel hadden dat we van tevoren gedacht zouden hebben en dat als we succes voor deze doelgroep zouden willen voorstellen we niet moeten schromen om te koersen op eigenschappen die we in het dagelijks leven misschien niet zo aangenaam zouden vinden. Tevens wil ik laten zien dat het onderzoeken van een doelgroep een opmaat voor een opleiding kan zijn.

In totaal heb ik 20 bekende televisie-Nederlanders getest. Ik werd benaderd eerst door BNN en later door Yorin met de vraag of ik bekende Nederlanders aan een test wilde onderwerpen. Dat leek me geweldig interessant. Voor een bedrijfspsycholoog is het bijzonder boeiend om mensen die bekend zijn van de televisie eens aan je testbatterij te mogen onderwerpen. Deze 20 mensen waren vaak geziene televisie-presentatoren, bekende modellen, beroemde zangeressen, zeer succesvolle sporters (winnaars van olympische medailles) en voor de liefhebber ook zeer bekende soapsterren. Het was opvallend dat deze mensen een behoorlijk gemeenschappelijk profiel hadden. In het onderstaande geef ik de groepsresultaten weer, en vermeld daar meteen bij welke eisen deze eigenschappen stellen aan een eventuele opleiding tot Bekende Nederlander. Daarbij moet wel de kanttekening gemaakt worden dat de aantallen wetenschappelijk gezien geen harde conclusies rechtvaardigen, maar ons denken wel een eind op weg kunnen helpen.

Zeer sociabel, bijzonder communicatief

BN'ers zijn niet gewoon bovengemiddeld extravert, maar maximaal extravert. Dus heel open, heel sociabel, heel joviaal. Ze zijn heel makkelijk pratend en ook zeer nadrukkelijk daardoor aanwezig. Hun communicatief vermogen spat er van af!

EISEN AAN DE OPLEIDING

Bied een scala aan mogelijkheden om in deze communicatiebehoefte te voorzien. Draag er zorg voor dat het onderwijs veel interactie biedt, hanteer veel onderwijsvormen waar de student gevraagd wordt iets te presenteren, te oefenen in de praktijk. Verder is het inhoudelijk van belang dat de student maximaal geprikkeld wordt om te leren de show te stelen en opdrachten te doen waarbij opvallen een *conditio sine qua non* is!

Zeer solistisch, weinig gericht op samenwerking

BN'ers zijn mensen die vooral met zichzelf bezig zijn, weinig over hebben voor een ander, egocentrisch zijn en ook behoorlijk scherp kunnen zijn als iets ze niet

bevalt. Uitermate assertief. Maar hoewel charmant, zijn ze over het algemeen niet zo vriendelijk.

EISEN AAN DE OPLEIDING

Hoewel er veel mogelijkheden moeten zijn om te communiceren met anderen, moet het ook zo zijn dat de opleiding veel mogelijkheden biedt om individueel te excelleren, misschien wel voor een echt tv-publiek. Vermijd al te veel teamopdrachten, en leer ze liever hoe al te veel samenwerking de persoonlijke carrière wel eens in de weg kan staan. Leer ze egocentrisch handelen, leer ze eenzaam zijn. Leer ze van zich af te bijten. Assertiviteit is een noodzakelijke overlevingscompetentie.

Chaotisch

Het zijn extreem slordige mensen. Met uitzondering van de sporters, zijn ze weinig planmatig, laten het leven op zich afkomen, houden niet van vooruitdenken. Ze zijn vooral intuïtief.

EISEN AAN DE OPLEIDING

De opleiding moet vrijheid bieden in het programma om tegemoet te komen aan het gebrek aan discipline, maar moet tegelijkertijd hulp bieden om het al te chaotische karakter op te vangen. Wellicht is ondersteuning van studenten bedrijfskunde een oplossing. Te gestructureerde studenten moet dat snel worden afgeleerd, want planmatigheid is geen succesfactor

Keihard werkend

Ze hebben zeer veel energie en zijn zeer hardwerkend. Stellen zichzelf en anderen bijzonder hoge eisen. Niet makkelijk voor zichzelf, maar ook niet voor anderen.

EISEN AAN DE OPLEIDING

Als de opleiding voldoet aan de overige eisen, is dit punt iets wat zijn positieve werk vanzelf zal doen. De kunst is om de grote gedrevenheid te vangen in de structuur van de opleiding. Eén aspect van het omgaan met hun grote energie kan wel een opleidingselement zijn, namelijk besteed je het aan zo lang mogelijk feesten of aan de uitbouw van je carrière?

Grote ego's, vaak overdreven zelfvertrouwen en heel gevoelig voor kritiek

Bekende Nederlanders laten hele dubbele eigenschappen zien als het gaat om hun stabiliteit. Ze hebben wel heel veel zelfvertrouwen en zelfwaardering, op basis waarvan ze zichzelf durven laten zien, maar ze

kennen ook een hele grote gevoeligheid voor kritiek. Ze zijn daar heel sterk mee bezig, en zijn er ook sterk door van slag als de kritieken negatief zijn.

EISEN AAN DE OPLEIDING

Voor iemand die bekend wordt (daar gaan we in deze fictieve opleiding van uit) en die dagelijks op straat aangesproken wordt, waarvoor mensen op straat zelfs flauwvallen (!), die in de krant komt, die bekritiseerd wordt, die – zeker als hij of zij jong is – geleefd wordt, is het zeer belangrijk dat deze naast gevoeligheid voor signalen uit de omgeving ook leert om op zichzelf te vertrouwen. Met andere woorden: een zeer belangrijk vak in zo'n opleiding zou het vak 'Personality building' moeten zijn, leren omgaan met kritiek, met het serieus nemen van je eigen intuïtie, van je eigen innerlijke koers. Ook om goed om te gaan met de verleidingen die de wereld van het succes biedt (geef voorlichting over drank, drugs, enzovoort). Ik moet bij deze laatste ontwerpaanbeveling wel een kanttekening plaatsen. Toen ik deze gedachten op een lezing eerder uitte, kwam de opmerking dat je op een opleiding voor deze mensen ze niet moest leren zeker te worden, maar dat je ze juist moet leren om onzeker te zijn. Daar zit zeker een kern van waarheid in. En deze waarheid werd voor mij onlangs bevestigd toen ik de bevindingen hoorde van een studente die op een toneelschool opgeleid was. Zij vertelde uitvoerig hoe mensen daar als het ware werden afgebroken om daarna weer te worden opgebouwd, gemodelleerd naar de inzichten van de docenten. Dat sluit erg aan op de gedachte dat je te zekere mensen, wil je ze succesvol maken in het metier van bekende Nederlander, meer onzeker moet maken als ze in dat beroep succesvol willen worden.

Creatief, veel ideeën

Deze mensen zijn zeer autonoom, hebben een grote verbeeldingskracht, houden van de vernieuwing, zijn erg creatief.

EISEN AAN DE OPLEIDING

De opleiding bevat veel vrije opdrachten om de creativiteit ruimte te geven. Het 'out-of-the-box denken' is een hoofdvak.

Wie bovenstaande typeringen enigszins meewarig leest en er zo het zijne van denkt: vergeet niet dat deze mensen succesvol zijn! Dat de kenmerken voor ons niet-bekende Nederlanders wellicht wat grotesk of vreemd overkomen wil niet zeggen dat ze daarmee niet waar zijn. Dat zijn ze juist wel. Succes komt nu eenmaal soms

in vreemde gedaanten. Om succes van deze bekende Nederlanders te voorspellen, moeten we dus in hoge mate afstand doen van wat we sociaal wenselijk vinden of wat we zelf in onze huis-, tuin- en keukensfeer prettig zouden vinden. De echt Bekende Nederlander moet een bijzonder moeilijk mens zijn om mee samen te wonen of samen te werken.

In het verlengde daarvan moet een opleiding ertoe bijdragen dat de studenten een aantal gedragingen leren die in het algemeen als onaangenaam zullen worden beoordeeld. Echter, zonder dat gedrag geen succes. Als de opleiding dit gedrag niet stimuleert of leert, is het derhalve geen goede opleiding. Dat we met de leerkrachten enig medelijden zullen krijgen, mag wel verwacht worden!

Assessment en opleidingsontwerpen

Als we ons verdiepen in de kenmerken van deze doelgroep, kom ik enigszins brainstormend tot bovenstaande kenmerken waar een opleiding aan zou kunnen voldoen. Maar zo zouden de ontwerpers van opleidingen deze ook kunnen enten op geheel andere doelgroepen op basis van de groepsresultaten verkregen uit assessments.

In mijn praktijk leveren wij regelmatig een bijdrage aan management-developmentstrajecten. Door de resultaten te bundelen verkrijgen wij een goed en specifiek beeld van de sterke en zwakke eigenschappen van de doelgroep. Op basis hiervan kan een MD-programma op maat worden gemaakt. Een vergelijkbare interventie kan uiteraard ook in het onderwijs worden gerealiseerd. Door een grondige kennis van de doelgroep kan het onderwijsprogramma op maat worden gemaakt. Overigens is het van hetzelfde grote belang om eerst grondig te analyseren wat mensen in de praktijk succesvol maakt. Die succesfactoren afstemmen op de doelgroep is een uitdaging. Een uitdaging die naar ik hoop zeker haalbaar is, zoals ik in dit artikel onder andere heb proberen te laten zien.

VALKUILEN BIJ ASSESSMENT

Nadat we de toegevoegde waarde van het assessment hebben beargumenteerd en ons verdiept hebben in factoren die spelen bij het voorspellen van succes, is het nu zaak om de (aspirant-) professionele beoordeelaar bij de hand te nemen en hem of haar te wijzen op enkele valkuilen die spelen bij de beoefening van het vak.

Voorspellen is een aardige bezigheid, die een verslavende kant kent en waarbij de voorspeller zich, als hij niet oppast, een beetje verheven kan gaan voelen. Van de volgende elementen merk ik dat ze dagelijks goed in de gaten moeten worden gehouden om het werk goed te blijven doen. Het zijn drie valkuilen waarmee

de toepasser van het assessment rekening moet houden, als hij zijn verantwoordelijkheid als professional serieus neemt.

De eerste valkuil: de waarheid van het te algemene woord

Eén valkuil wordt geïllustreerd door het volgende voorbeeld dat ik aangereikt heb gekregen van professor Bonarius, voormalig hoogleraar persoonlijkheidsleer aan de Rijksuniversiteit Utrecht.

Een onderzoeker doet het volgende experiment: hij spreekt aan de universiteit een groep studenten toe en vraagt of ze mee willen doen aan de ontwikkeling van een nieuwe vragenlijst. Alle studenten, vanaf nu proefpersonen, vullen de vragenlijst in en krijgen na een maand de uitslag thuisgestuurd. Deze uitslag is netjes opgesteld, geeft blijk van gedegenheid, is persoonlijk, en uiteindelijk wordt gevraagd of ze een enquête willen invullen. Deze enquête is eenvoudig en stelt de vraag of mensen zich herkennen in de uitslag. De uitslag was overtuigend; het merendeel herkende zich heel behoorlijk. Nou zult u zeggen, dat is dan mooi en het is dus een goede test. Wellicht, maar het venijn zit in de staart van dit verhaal, want het experiment bestond eruit dat alle proefpersonen dezelfde persoonsbeschrijving thuis hadden gekregen. Ziedaar het effect van een opgestelde beschrijving die in algemene termen is en weinig concreet. Dit effect, ik noem het maar het Libelle-astrologie-effect (waarom denkt u dat u zich herkent in de astrologiebeschrijvingen?), maakt ons beoordelaars lui. De kandidaat herkent zich eigenlijk te snel in hetgeen over hem of haar is opgeschreven. Ik zal u een voorbeeld geven waarin u zich vast herkent.

Het altijd goed herkenbare rapport:

U neemt besluiten op momenten dat ze u goed uitkomen en toont zich in de communicatie behoorlijk open, maar laat niet te vaak het achterste van uw tong zien. In situaties waar u leiding moet geven probeert u rekening te houden met anderen, behalve als het echt niet anders kan, dan kiest u voor het doel. U werkt hard, heeft behoefte aan tijdige ontspanning en kent uw gevoelens van onzekerheid, maar weet deze over het algemeen goed te controleren, u heeft oog voor de grote lijn, maar minder affiniteit met de dagelijkse details, bla, bla, bla.

Dit soort teksten roepen over het algemeen een hoge mate van herkenning op, maar zijn vergelijkbaar met uitspraken als: het wordt in Nederland iedere dag licht en natuurrampen daargelaten ook iedere dag weer donker. Met andere woorden: ze zijn altijd waar, maar hebben de schijn van specificiteit. Dit effect wordt versterkt door het gegeven dat in mijn praktijk de deelnemer ook bewust of onbewust zal denken dat het wel goed zal zijn, omdat ik het heb opgeschreven, een

ervaren adviseur van een gerespecteerd bureau. Om te komen tot een zo specifiek mogelijke uitspraak moet u de competenties die relevant zijn voor de functie scherp in kaart brengen, daar goede valide en betrouwbare meetinstrumenten op loslaten en vervolgens zo specifiek mogelijk uitspraken over doen, waarbij wel duidelijk is wat de kans is op het voorkomende gedrag.

De tweede valkuil: het doen van absolute uitspraken

Wat verder belangrijk is is dat de voorspeller zich realiseert dat hij altijd praat in termen van waarschijnlijkheid, van kansen. Voorspellen in de sociale wetenschap (niet alleen in de sociale wetenschap) is behoorlijk moeilijk als het gaat om het individu. Je moet voorzichtig zijn, je mag nooit een te absolute uitspraak doen. Wie in assessments doet, die doet dan ook in

Wie in assessments doet, die doet dan ook in kansen

kansen. Want je zult maar zeggen tegen iemand dat hij/zij niet geschikt is voor een functie of beter aan een andere loopbaan kan denken en dat achteraf blijkt dat dat in zeven van de tien gevallen wel klopt, maar dat deze meneer of mevrouw nou net bij die drie hoort. Het zal overigens duidelijk zijn dat valkuil 1 en 2 het de beoordelaar niet makkelijk maken, omdat een specifieke uitspraak het risico creëert van een te absolute uitspraak. Omgekeerd is een uitspraak waarbij je teveel wikt en weegt voor een opdrachtgever weer onduidelijk, want wat moet deze nou? Iemand aannemen of juist niet? Die opdrachtgever wil juist weer wel de absolute uitspraak. Het is voor de psycholoog of de beoordelaar balanceren. Het goed omgaan met deze balans vraagt een hoge mate van professionaliteit.

De derde valkuil: na een goede diagnose is de beoordelaar klaar

Een belangrijke ervaring is voor mij de hiernavolgende. Het leidt tot een relativering van wat het meten is en wat het belang van opleiding en begeleiding onderstreept en daarmee hetgeen eerder is gesteld, namelijk dat assessments onderdeel moeten zijn van een goed geïntegreerd HRM-beleid.

Het voorbeeld dat ik enkele jaren geleden hoorde was het volgende. Wie in Nederland een eigen bedrijf start, kan vanuit verschillende instanties ondersteuning krijgen. Ik sprak op een goed moment met zo'n begeleider. Deze vertelde het volgende: het landelijk sla-

gingspercentage van startende ondernemers was bijvoorbeeld 20 procent en zijn slagingspercentage was ruim het dubbele, en dat al vele jaren lang. Hij had niet een andere doelgroep, hij had ook geen andere slagingsnormen. Het enige wat hem nadrukkelijk onderscheidde van zijn collegae was een Emile Ratelband-instelling van positief denken. Hij wist zijn starters zo te empoweren, ze zo te motiveren, dat deze veel beter bestand waren tegen tegenslag en veel langer doorgingen dan andere starters. Dat maakte hen duidelijk succesvoller dan anderen. Ik denk niet dat als we deze mensen aan een assessment hadden onderworpen er een significant betere groep naar voren was gekomen. Maar de kwaliteit van de begeleiding was hier dermate goed dat deze een wezenlijke invloed had op het resultaat, op het succes.

Met deze drie valkuilen voor ogen zal het duidelijk zijn dat het doen van assessments niet altijd even makkelijk is en van degene die het toepast een grote mate van bewustzijn vraagt van de reikwijdte van zijn uitspraken. De mogelijkheid bestaat van te algemeen en daarmee nietszeggend (zal vaak niet gewaardeerd worden door de opdrachtgever), de mogelijkheid bestaat van te absolute uitspraken en daarmee bijvoorbeeld anderen, kinderen, studenten stigmatiseert en afremt (het zal duidelijk zijn dat de deelnemer dat niet plezierig zal vinden) en het risico bestaat dat u achterover gaat leunen na de diagnose en vergeet dat te doen wat het belangrijkste is, namelijk medewerkers/studenten motiveren en enthousiasmeren en zelfvertrouwen geven om belangrijke stappen in hun leven te zetten. Het assessment biedt veel meer inzicht in de mogelijkheden van groepen studenten, waardoor bijvoorbeeld bedrijven, scholen, universiteiten veel beter in staat zijn om hun onderwijs/trainingsinspanningen te richten op mogelijkheden van specifieke doelgroepen. Tevens biedt het werknemers, studenten de unieke mogelijkheid om vroeg in hun leven meer objectieve informatie te verkrijgen over hoe de vlag er nu bij hangt, kijkend naar hun eigen competenties. Zij hebben dan soms al jong de mogelijkheid om gericht te werken aan hun toekomst.

TOT SLOT

In dit artikel heb ik willen illustreren dat het toepassen van assessments een zaak is van verantwoordelijkheid nemen. De opdrachtgever neemt verantwoordelijkheid doordat hij middelen zorgvuldig besteedt en mensen zorgvuldig behandelt. De deelnemer aan het assessment neemt verantwoordelijkheid door zich bewust te zijn van de organisatie- en persoonlijke belangen die er spelen. En bovenal de toepasser van het assessment voelt zich verantwoordelijk voor de grenzen, maar ook de mogelijkheden die zijn resulta-

ten hem opleggen, c.q. bieden. Dat deze grenzen hem of haar voor lastige uitdagingen stellen, maakt het vak van beoordelaar interessant, dat de mogelijkheden nog veel groter zijn dan we met zijn allen vaak denken, maakt het alleen maar nog boeiender. Een assessment is niet de panacee voor alles, maar het is wel een uitermate vruchtbare HRM-activiteit, waarmee de organisatie de aandacht voor de medewerker, of in het onderwijs de student, vergroot en zijn of haar kansen op gepast succes eveneens. En wie wil dat niet?!



Literatuur

- Damen, B. en A.C. van der Horst, *Inzicht in commercieel succes*. Breda: GTP uitgave, 1992.
- Drenth, P.J.D., *Inleiding in de testtheorie*. Deventer: Van Loghum Slaterus, 1985.
- Hoekstra, H.A. en E. van Sluis, *Management van competenties*. Assen: Van Gorcum, 1999.

$$\alpha^2 + \beta^2 = \gamma^2$$



HET DESIGNPROCES VAN ASSESSMENT CENTERS:

SCHIPPEREN TUSSEN PRAKTIJK EN KWALITEIT

FILIP LIEVENS EN NANJA KOLK

In een assessment center worden belangrijke beslissingen genomen op basis van iemands prestatie. Vanzelfsprekend was en is de kwaliteit van deze beslissingen onderwerp van talloze studies. Echter, de aanbevelingen die hieruit volgen, moeten in de praktijk ook implementeerbaar zijn. In dit artikel belichten we deze twee perspectieven met als doel de kloof tussen wetenschap en praktijk te overbruggen.

INTRODUCTIE

Vanaf de jaren tachtig werd het assessment center gepositioneerd als hét instrument dat zowel prestatie van mensen kan voorspellen als een betrouwbare indicatie kan geven van doorgroeipotentieel. Het assessment center is bij uitstek geschikt voor het selecteren van managers, omdat het recht doet aan de complexiteit en multifunctionaliteit van hun functie. In een assessment center nemen kandidaten deel aan verschillende (simulatie)opdrachten en testen, die tezamen de beoogde managementfunctie representeren. Voorbeelden zijn het tweegesprek (bijvoorbeeld met een veeleisende klant of een onwelwillende werknemer), een groepsdiscussie (wie ontpopt zich als de leider?) en een postbakoefening (hoe baant iemand zich een weg door stapels memo's, brieven en e-mails?). Deze opdrachten worden vaak gecombineerd met intelligentietesten en (persoonlijkheids)vragenlijsten.

Verskillende assessoren (beoordelaars) beoordelen de prestaties van deelnemers op een aantal relevante competenties. Na afloop van alle opdrachten (met een tijdsduur van één tot twee dagen) volgt er een consensusdiscussie (assessorenvergadering) die leidt tot een uitspraak over iemands kans van slagen in een baan (selectie), of iemands potentieel in een ontwikkelingstraject (development center).

De doelstelling van dit artikel is tweevoudig:

- Allereerst beogen wij de diverse problemen en knelpunten bij het design van assessment centers in kaart te brengen. Deze problemen worden bekeken vanuit twee perspectieven: een wetenschappelijk perspectief, waarbij de nadruk ligt op het waarborgen van de *kwaliteit* van het assessment center, en een pragmatisch perspectief, waarbij het vooral gaat om *knelpunten bij de implementatie* van het assessment center.
- De tweede doelstelling van dit artikel is het aandragen van mogelijke oplossingen voor de geïdentificeerde problemen en knelpunten. Bij het bespreken van de problemen van het assessment center, alsook de door ons aangedragen oplossingen, gebruiken we het assessment-centerdesign als kapstok. Achtereenvolgens bespreken we daarna de problemen en oplossingen binnen respectievelijk het kwaliteitsperspectief en het pragmatische perspectief. We zullen

afsluiten met een overzicht van kwaliteitsrichtlijnen en aanbevelingen voor verbetering van het assessment center.

ASSESSMENT CENTERS BINNEN HRM

Assessment centers zijn vooral bekend als instrument om een momentopname te maken voor het nemen van plaatsingsbeslissingen. Het is inderdaad zo dat bij de *eerste generatie* assessment centers het accent lag op het bepalen van de geschiktheid van de kandidaat. Er werd weinig tot geen tijd gespendeerd aan de ontwikkeling van de kandidaat (Thornton, 1992). De doelstellingen van een assessment center zijn in de loop van de jaren aanzienlijk verruimd, waardoor assessment centers een meer centrale rol zijn en zullen gaan spelen binnen HRM.

Het feit dat assessment centers breder kunnen worden ingezet dan voor selectiedoelen blijkt bijvoorbeeld uit de opkomst van de *tweede generatie* assessment centers. Dit zijn de zogenoemde development centers of ontwikkelingsgerichte assessment centers. Een ontwikkelingsgericht assessment center (bijvoorbeeld in het kader van managementontwikkeling of interne doorstroming) heeft een 'diagnostische functie', waarbij de sterke en zwakke punten van managers, op basis van de door het assessment center gemeten gedragskenmerken, in kaart worden gebracht. De assessoren zijn hierbij niet enkel beoordelaars, maar fungeren ook als coaches en facilitatoren. Op deze manier komen mogelijke ontwikkelingsbehoeften van de deelnemers aan de oppervlakte. Zo kunnen verantwoorde trainings- en ontwikkelingsprogramma's worden ontwikkeld om de zwakke plekken bij te spijkeren.

Bij de *derde generatie* assessment centers krijgen de deelnemers tijdens het assessment center na elke oefening feedback, zodat de deelnemers worden aangemoedigd om te gebruiken wat ze in de vorige oefening hebben geleerd. Er ligt dus een heel sterk accent op training. De assessment-centersimulaties (bijvoorbeeld vergadering, functioneringsgesprek) worden aangewend als oefeningen om managementvaardigheden aan te leren en/of bij te spijkeren. De assessoren fungeren zuiver als coaches en facilitatoren. Het plannen van ontwikkeling krijgt aandacht tijdens maar ook na het assessment center, en deelnemers worden actief betrokken bij het genereren van hun assessment.

Daarnaast zien we steeds vaker dat assessment centers worden gebruikt als een methodiek om de organisationele en individuele doelen op elkaar af te stemmen. Zo bereikt men een geïntegreerd HR-beleid (Ballantyne en Povah, 1995; Seegers, 1997; Steltenpohl, 1993) dat afgestemd is op de visie/strategie van de organisa-

tie. De kern hierbij is dat menselijke prestaties als centrale component binnen de HR-cyclus (Fombrun, Tichy en Devanna, 1984) gebaseerd zijn op 'gedrag' en dat het management een gemeenschappelijke taal krijgt om dit (arbeids)gedrag te bespreken. Als deze prestatie gedefinieerd en vertaald wordt aan de hand van gedragskenmerken zoals die in een assessment center worden gemeten, dan is het assessment center een bindend element tussen de verschillende componenten van deze cyclus, namelijk selectie, beoordeling, ontwikkeling en beloning. Op deze manier fungeert gedrag, zoals benaderd in een assessment center, als uitgangspunt om de verschillende en vaak los van elkaar staande HR-activiteiten op elkaar af te stemmen en te komen tot een meer geïntegreerd HR-beleid (Ballantyne en Povah, 1995).

HET DESIGNPROCES VAN ASSESSMENT CENTERS

In het Nederlandstalige taalgebied waren Seegers en Esser (1982) een van de eersten die een tekst publiceerden over het maken van een design van een assessment center. Nog steeds geldt het model van Seegers en Esser als een mogelijke blauwdruk voor het designproces van assessment centers. Zij onderscheidden vier designfasen: inventarisatie, operationalisatie, uitvoering en evaluatie. Elke fase omvat een aantal specifieke aspecten waaraan een organisatie nauwgezet aandacht moet schenken. Figuur 1 geeft een schematisch overzicht van deze fasen en de belangrijkste aandachtspunten.

Inventarisatiefase

In de inventarisatiefase staan drie kernvragen centraal. Eerst moet de doelstelling van het assessment center worden bepaald. Assessment centers kunnen immers voor selectie, diagnose en ontwikkeling worden gebruikt. Vervolgens dient de ontwikkelaar de doelgroep af te grenzen. Met andere woorden: voor welke functies wordt het assessment center ontwikkeld? Meestal betreft het managementfuncties. Ten derde worden de inhoud en de vereiste competenties (i.e. vereiste kennis, vaardigheden en eigenschappen) van de gekozen functie(s) in kaart gebracht (bijvoorbeeld planning en organiseren, oordeelsvorming, besluitvaardigheid).

Operationalisatiefase

De tweede fase in de ontwikkeling van een assessment center impliceert de operationalisatie van de middelen die nodig zijn voor de uitvoering (assessoren, assesses, opdrachten en beoordelingsformulieren). Hierbij vormt de selectie van assessoren een eerste

belangrijk aandachtspunt. Hiervoor wordt een beroep gedaan op psychologen, op personeelsleden van de organisatie of op externe adviseurs. De training van deze assessoren begint met informatie over assessment centers, de gehanteerde competenties en de gebruikte opdrachten. Daarnaast is het de bedoeling dat assessoren op een praktijkgerichte manier vaardigheden leren, zoals het observeren en systematisch beoordelen, het integreren van de verschillende beoordelingen in een rapport en het geven van feedback aan kandidaten.

De selectie van assessee's vormt een tweede actiegebied binnen de operationalisatiefase. Het aantal kandidaten per assessmentsessie bedraagt meestal vier tot acht zorgvuldig gekozen personen. Met zorgvuldig bedoelen we dat de assessee's vaak al een uitgebreide screening achter de rug hebben voordat ze worden toegelaten tot het assessment center. Belangrijk zijn ook de instructies die de kandidaten krijgen. Zo dienen de kandidaten ingelicht te worden over de inhoud en de doelstellingen van het assessment center. Ook is het zo dat ze dienen te weten welke gegevens over hen (niet) worden bijgehouden en wie er inzage in heeft.

De selectie van opdrachten (bijvoorbeeld de postbak, het tweegesprek) vormt het derde en laatste domein binnen de operationalisatie van een assessment center. De geselecteerde testen en opdrachten zijn een functie van de doelstellingen en de competenties van het assessment center. Dat wil zeggen dat de doelstellingen en de geselecteerde competenties bepalen welke opdrachten zullen worden opgenomen.

Uitvoeringsfase

Als men beschikt over beoordelaars, een groep kandidaten en een verzameling testen en opdrachten, kan worden overgegaan tot de feitelijke uitvoering van het assessment center. Assessoren observeren en beoordelen de kandidaten in de diverse opdrachten. Hierbij kunnen de assessoren eventueel gebruikmaken van observatiechecklists, beoordelingsschalen en video-opnames. Tijdens iedere simulatie dienen zij in elk geval een verslag van de observaties te maken. Zij nemen in dit verslag beoordelingen van de gedragscriteria op, gebaseerd op relevante gedragingen. Een goede planning en organisatie van dit assessment-centerprogramma zijn essentieel, omdat de assessoren en kandidaten vlot moeten kunnen roteren tussen de verschillende opdrachten.

Evaluatiefase

Na het assessment center dient volgens Seegers en Esser (1982) een evaluatie van het assessment center plaats te vinden. In eerste instantie gebeurt deze evaluatie door de assessoren zelf. Zij brengen immers de

informatie over de kandidaten uit verschillende opdrachten bij elkaar om tot een globaal oordeel te komen. Het assessment center kan ook op een meer systematische manier geëvalueerd worden. Dit kan bijvoorbeeld een studie naar de voorspellende kracht of de billijkheid van het assessment center zijn. Ook kan een informele rondvraag bij kandidaten en assessoren over de sterke en zwakke punten van het georganiseerde assessment center plaatsvinden.

Fase 1 – Inventarisatie

- Doelstelling van het assessment center
- Functies waarvoor het assessment center bedoeld is
- Functiebeschrijving en -analyse

Fase 2 – Operationalisatie

- Selectie en training van assessoren
- Selectie en instructie van kandidaten
- Selectie en ontwikkeling van oefeningen

Fase 3 – Uitvoering

- Samenstelling draaiboek (planning, rotatie assessoren, et cetera)
- Daadwerkelijk assessment center (duur, plaats, et cetera)

Fase 4 – Eindevaluatie

- Eindoordeel en rapportage
- Feedback naar kandidaat en onderneming
- Evaluatie van assessment-centerprogramma

Figuur 1. Gefaseerd model van het design van assessment centers (naar Seegers en Esser, 1982).

PROBLEMEN EN OPLOSSINGEN BINNEN HET KWALITEITSPERSPECTIEF

Binnen de designfasen van Seegers en Esser (1982) doen zich verschillende problemen voor die de kwaliteit van het assessment center kunnen schaden. Aangezien er belangrijke beslissingen worden genomen op basis van iemands prestatie in een assessment center, was en is de kwaliteit van deze beslissingen onderwerp van talloze empirische studies. In deze wetenschappelijke studies zijn niet alleen de problemen geïdentificeerd, maar zijn ook een aantal bruikbare suggesties gedaan om de kwaliteit te verbeteren (voor overzichten hieromtrent zie Kolk, 2001; Lievens, 2002; Lievens en Conway, 2001).

Inventarisatiefase

In de inventarisatiefase is het *selecteren van de juiste hoeveelheid competenties* problematisch. Het is zeer aantrekkelijk, omwille van efficiency, om zo veel mogelijk com-

petenties per opdracht te proberen te meten. Onderzoek laat echter systematisch zien dat assessoren niet in staat zijn meer dan drie tot vijf competenties te onderscheiden binnen één opdracht.

Ten tweede blijkt het *combineren van de juiste competenties binnen één opdracht* niet eenvoudig. Als men bijvoorbeeld de competenties 'sensitiviteit' en 'luisteren' in één opdracht wil meten, dan zal het gedrag dat bij deze competenties hoort in de ogen van de assessoren zó op elkaar lijken, dat de beoordelingen nagenoeg hetzelfde zullen zijn. Om sensitief te kunnen zijn, moet men immers ook goed kunnen luisteren. Met andere woorden: dan kan men net zo goed alleen de competentie 'sensitiviteit' opnemen. Het gebruik van conceptueel onafhankelijke competenties binnen dezelfde opdracht zorgt dus voor een beter onderscheidend vermogen en verhoogt daarom de kwaliteit van de beoordeling.

Een praktische suggestie binnen dit probleemgebied richt zich op het gebruik van drie domeinen of categorieën van competenties: denken, voelen en kracht. Wanneer men in een opdracht steeds één competentie uit elk van de domeinen selecteert (bijvoorbeeld 'luisteren' uit het domein 'voelen', 'inzicht' uit het domein 'denken' en 'doorzettingsvermogen' uit het domein 'kracht'), dan blijkt de kwaliteit van de beoordelingen substantieel te verbeteren (Kolk, 2001).

Ook het *selecteren van de juiste opdrachten* is van belang. De opdrachten worden idealiter op maat voor de organisatie of functie ontwikkeld. Zij dienen daarom gebaseerd te zijn op een competentieanalyse. Assessment-centeropdrachten zijn daarentegen geen job samples (arbeidsproeven), zoals een typetest voor typisten. Vooral bij complexe functies, zoals vele managementfuncties, is het niet zinvol te trachten de functie letterlijk te simuleren.

De mate waarin een opdracht gelijkenissen vertoont met de werksituatie is belangrijk omdat het de acceptatie van de assessment-centeruitkomsten vergroot. Indien een opdracht echter te veel op de werksituatie

Kandidaten moeten in iedere opdracht de kans krijgen gedrag te laten zien dat gerelateerd is aan de te meten competentie. Opdrachten dienen daarom specifiek ontwikkeld te worden om competentiegerelateerd gedrag uit te lokken. Bijvoorbeeld, als sensitiviteit een belangrijke competentie is, dan moet de opdracht veelvuldig de kans bieden sensitief gedrag te laten zien.

Operationalisatiefase

Wanneer het assessment center operationeel is, dan blijken er zich problemen voor te doen die te maken hebben met *beoordelaarseffecten* bij assessoren, zoals strengheid, mildheid, centrale tendentie of halo. Met name halo verslechtert de kwaliteit van de beoordelingen. Wanneer er halo optreedt in de beoordelingen, betekent dit dat verschillende competenties zo sterk met elkaar samenhangen dat er in feite geen onderscheid meer wordt gemaakt. Er kan halo optreden als de betekenis van de competentie overlapt (zoals in het voorbeeld van 'sensitiviteit' en 'luisteren'). Dit kan voorkomen worden door conceptueel onafhankelijke competenties te selecteren. Maar halo treedt ook op door een foutieve denkwijze van assessoren. Beoordelaars vormen zich bijvoorbeeld gedurende een opdracht een algemene indruk van de kandidaat, die ze laten weerklinken in alle competentiescores (*general impression halo*). Een ander voorbeeld van halo is dat beoordelaars één competentie het meest relevant vinden, en alle competentiescores baseren op de prestatie van de kandidaat op die ene competentie (*salient feature halo*).

Om zo min mogelijk beoordelaarsfouten te maken, dienen assessoren zo veel mogelijk ervaring te hebben met de beoordelingstaak. Assessoren dienen daarom ruimschoots de tijd te nemen om ervaring op te doen in de beoordelingstaak, vooraleer de beoordelingen worden gebruikt om beslissingen te nemen. Het is verder raadzaam om zowel (P&O-)managers als psychologen in het assessorenteam te betrekken, omdat beide assessortypen een verschillend aspect van werkprestatie voorspellen.

De kwaliteit van het assessment center verbetert bovendien wanneer assessoren getraind worden in het beoordelen volgens een frame-of-reference training (Lievens, 2001). Dit type training voorziet assessoren van een gemeenschappelijke 'prestatietheorie', zodanig dat elke assessor de prestatie van kandidaten beoordeelt op basis van dezelfde definitie van wat effectieve prestatie precies inhoudt.

Naast de beoordelaar is ook de *rolspeler* van belang. Het is belangrijk dat de rolspelers, of dat nu psychologen of acteurs zijn, de kandidaten steeds *volgens één bepaalde methode tegenspielen*. Er moet een duidelijke leidraad

Het gebruik van conceptueel onafhankelijke competenties verhoogt de kwaliteit van de beoordeling

lijkt, zijn kandidaten met ervaring in die specifieke werksituatie in het voordeel ten opzichte van kandidaten zonder deze ervaring. Men zou daarom naar een compromis kunnen zoeken tussen een replica van een functie en een opdracht die ongerelateerd is aan die functie. Opdrachten zouden bijvoorbeeld een bepaalde industrie kunnen representeren, maar niet de exacte functie of organisatie.

zijn: doet de kandidaat zus, dan is de reactie zo. De gedragingen van de rolspeler dienen op die manier als ‘platforms’ voor de reacties van de kandidaat. Deze platforms moeten voor zowel rolspelers als assessoren duidelijk zijn. Platforms zorgen ervoor dat het juiste gedrag wordt uitgelokt (gedrag dat gerelateerd is aan de competentie) en bovendien dat er gevarieerd gedrag wordt uitgelokt (zodanig dat alle competenties aan bod komen). Hierdoor kunnen assessoren daadwerkelijk zien of een kandidaat ook echt die verschillende competenties in huis heeft.

Uitvoeringsfase

Onderzoek laat zien dat het raadzaam is assessoren te laten roteren na iedere opdracht, zodat iedere assessor elke kandidaat slechts één keer beoordeelt (Kolk, Born en Van der Flier, in druk). Hierdoor geven de opdrachten ieder een onafhankelijk beeld van de prestatie van kandidaten op bepaalde competenties, hetgeen de *betrouwbaarheid* van de competentiemeting verhoogt. Wanneer een assessor bij diverse opdrachten aanwezig is, zouden bovendien contrasteffecten kunnen optreden: een assessor kijkt niet meer onbevooroordeeld naar een tweede opdracht wanneer een kandidaat in de eerste opdracht bijzonder slecht of juist goed heeft gepresteerd.

Onderzoek wijst verder uit dat er ook aan de *assessmentprocedure* veel te verbeteren valt. Zo verhoogt het beoordelen met behulp van checklists de kwaliteit van de beoordelingen (Reilly, Henry en Smither, 1990). Wel loopt men door het gebruik van checklists het risico dat assessoren teveel gericht zijn op het gedrag dat op de checklist voorkomt en voorbijgaan aan ander relevant gedrag. Toch is een checklist aan te bevelen bij onervaren assessoren zoals managers.

Een andere suggestie voor verbetering van de assessmentprocedure is het onthullen van de competenties aan kandidaten (een transparant assessment center). Het idee hierbij is dat deelnemers voorafgaand aan een assessment een bepaalde verwachting hebben van welk gedrag gewenst is. Dit idee kan juist of onjuist zijn. Wanneer de verwachting van gewenst gedrag onjuist is, laat de deelnemer gedrag zien dat ongerelateerd is aan de competenties die de assessoren beoordelen. Dit wil niet zeggen dat de deelnemer de gewenste competentie niet in huis heeft. Het onthullen van de competenties en de definities daarvan voorafgaand aan een assessment, verhoogt eenvoudigweg de kans dat de deelnemer gedrag laat zien dat gerelateerd is aan die competentie, en verlaagt de kans op het vertonen van irrelevant gedrag. Dit is gunstig voor de kwaliteit van het assessment center, de billijkheid (iedereen krijgt een gelijke kans) en voor de acceptatie van de resultaten (Kleinmann, Kuptsch en Köller, 1996; Kolk, in druk).

Evaluatiefase

Het probleem dat zich in deze fase voordoet is dat de *evaluatie van assessment centers* (als het al gebeurt) *lastig met elkaar te vergelijken* zijn. Zoals eerder al werd opgemerkt, verschillen assessment centers vaak van elkaar op verschillende aspecten. Omdat deze verschillen niet altijd duidelijk zijn, is het ook lastig een assessment center te vergelijken met een assessment-center ‘benchmark’. In principe zouden daarom de validiteit en de betrouwbaarheid van het assessment center frequent kwantitatief en kwalitatief geëvalueerd moeten worden. De resultaten hiervan dienen te leiden tot verbeteringen in het assessment center.

PROBLEMEN EN OPLOSSINGEN BINNEN HET PRAKTIJKPERSPECTIEF

Het model van Seegers en Esser (1982) vormt in zeker opzicht een ideaal model. Het geeft de diverse fasen in de ontwikkeling en implementatie van assessment centers gestructureerd weer en beklemtoont de verschillende beslissingen die door ondernemingen en adviesbureaus genomen dienen te worden. In de praktijk zien we echter dat de fasen op uiteenlopende manieren worden geïmplementeerd, waarbij er in bepaalde fasen vaak concessies worden gedaan. Bovendien staat dit model niet stil bij de mogelijke problemen die ondernemingen in de praktijk kunnen ervaren bij de ontwikkeling en implementatie van assessment centers.

Het laatste decennium zijn er zowel in Nederland als Vlaanderen kwalitatieve studies (Lievens en Goemaere, 1999; Van Dam, Altink en Kok, 1992) uitgevoerd, die de praktische knelpunten bij het design van assessment centers hebben opgespoord en samengevat. Bij beide kwalitatieve studies fungeerde het model van assessment-centerdesign van Seegers en Esser (1982) als conceptueel raamwerk bij het onderzoek.

Van Dam et al. (1992) interviewden 25 Nederlandse gebruikers van assessment centers om de problemen die kunnen optreden bij het praktisch gebruik van assessment centers in kaart te brengen. Lievens en Goemaere (1999) voerden hetzelfde onderzoek in Vlaanderen uit. Zij namen semi-gestructureerde interviews over de toepassing van assessment centers af bij 23 personeelsfunctionarissen. Deze werden geselecteerd uit een representatieve steekproef van assessment-centergebruikers (voor meer details over het onderzoek zie Lievens en Goemaere, 2000). Van deze 23 ondervraagde ondernemingen ontwikkelden er 8 volledig zelfstandig hun assessment center. De overige 15 ondernemingen deden voor de ontwikkeling een beroep op een extern adviesbureau.

We geven hieronder een systematisch overzicht van problemen waarmee Nederlands en Vlaamse onder-

nemingen in de praktijk bij het design van assessment centers geconfronteerd worden. Merk op dat het hier gaat over ondernemingen en niet over adviesbureaus.

Inventarisatiefase

Uit de studies van Van Dam et al. (1992) en Lievens en Goemaere (1999, 2000) kwam naar voren dat practici *zelden een grondige competentieanalyse uitvoeren* voorafgaand aan een assessment center. Dit was het geval in zowel Nederland als Vlaanderen. Er werd veelal gewerkt met een standaardlijst van essentiële competenties uit een aantal sleutelfuncties.

De definitie van competenties bleef bij de meeste ondernemingen constant over de verschillende assessment centers heen. Deze manier van werken verklaarde de problemen die de assessoren tijdens de uitvoeringsfase ondervonden met de inhoud van de competenties. *De definitie uit de standaardlijst kwam niet steeds goed overeen met de werkelijke inhoud van de competentie voor een specifieke functie.* Men hanteerde verder *dikwijls te veel competenties* (soms meer dan 15).

Een laatste probleem in verband met de competenties was dat *niet elke competentie* (bijvoorbeeld 'stressbestendigheid') *steeds in een opdracht geoperationaliseerd kon worden.*

Operationalisatiefase

Een knelpunt binnen de operationalisatiefase, dat enkel in Vlaanderen waargenomen werd, betrof het gebruik van lijnfunctionarissen als assessoren. Hoewel dit een aantal saillante voordelen (i.e. in het licht van de organisatiecultuur en latere acceptatie) met zich meebracht, lieten deze studies ook de keerzijde van de medaille zien. Kennis van de functie creëerde een referentiekader, waardoor de *beoordeling door lijnfunctionarissen gekleurd* werd. Lijnfunctionarissen wogen de kandidaat af ten opzichte van de ideale functie-uitvoerder of ten opzichte van bestaande personen en situaties. Dit fnuikte de objectieve beoordeling van het specifieke gedrag dat de kandidaat vertoonde tijdens de opdrachten.

Een andere opmerkelijke vaststelling was dat een *grondige training van assessoren vaak werd verwaarloosd*. Sommige respondenten stelden dat een opleiding een te grote investering vereiste, zowel financieel als in tijd. Precies om die reden werkte een aantal bedrijven uitsluitend met externe consultants als assessoren. Andere respondenten zagen de behoefte aan een assessorenopleiding niet in. Men ging ervan uit dat de personen die als assessor optraden voldoende ervaring hadden in het on-the-job observeren van mensen. Indien een assessorentraining plaatsvond, beperkte deze zich tot informatie over assessment centers, een

uiteenzetting van de techniek van beoordelen en het inoefenen van het onderscheid tussen observatie en interpretatie.

Uitvoeringsfase

Binnen de uitvoeringsfase bleek met name het observeren van kandidaten een belangrijk knelpunt. Zelfs als een assessorenopleiding werd aangeboden, doken er problemen bij het observeren en beoordelen van gedrag op. Nederlandse en Vlaamse practici merkten op dat het dikwijls zeer lastig was om assessoren het *verschil tussen observatie en interpretatie* bij te brengen. *Assessoren bleken bij het registreren van feitelijk gedrag geregeld in algemeenheden of voortijdige interpretaties te vervallen.* Ook de classificatie van de geregistreerde gedragingen in de vooropgestelde competenties vormde in de Nederlandse praktijk een ernstig probleem. Ondanks de gestandaardiseerde definitie van de competenties bleken assessoren steeds terug te grijpen naar een eigen interpretatie. Bij de gezamenlijke discussie leidde dit vaak tot oeverloze discussies over de omschrijving en de inhoud van de competenties. Checklists voor gedragsobservatie zijn nog niet ingeburgerd. Slechts de helft van de Vlaamse ondernemingen maakte er gebruik van.

De twee studies naar het praktisch gebruik van assessment brachten centers ook een aantal praktische problemen bij de planning en uitvoering van het assessment center aan het licht. Deze hadden voornamelijk te maken met de zware inspanning die de organisatie van een assessment center vereiste. In de eerste plaats was het geen sinecure om op zoek te gaan naar een geschikte dag voor de uitvoering. Doordat het niet eenvoudig is om kandidaten en assessoren telkens op de afgesproken tijd en plaats te mobiliseren, toonden de studies ook aan dat het daadwerkelijke assessment center *zelden volgens planning en timing verloopt*.

Het was verder zo dat diverse ondernemingen slechts één assessment-centeropdracht (bijvoorbeeld een rolspel) gebruikten en deze toch de naam 'assessment center' toekenden, hoewel in de definitie van assessment centers het belang van het gebruik van 'meervoudige' beoordelingstechnieken opdrachten benadrukt wordt.

Evaluatiefase

Een laatste bevinding van de twee studies was dat zowel in Nederland als in Vlaanderen de *evaluatie van het assessment center beperkt bleef* tot een rondvraag naar de ervaringen van de verschillende deelnemers. Een grondige evaluatie van het hele assessment-centerprogramma (bijvoorbeeld een studie naar de betrouwbaarheid of de voorspellende validiteit) was bijzonder zeldzaam. Organisaties zagen het belang van dit eva-

luatieonderzoek niet in en hadden vaak bezwaren tegen de extra kosten inherent aan evaluatiestudie.

CONCLUSIE: AANBEVELINGEN EN KWALITEITSRICHTLIJNEN

Beoordeling vormt een belangrijke pijler binnen HRM. Selectie, prestatiebeoordeling, remuneratie, potentieelbeoordeling, detectie van ontwikkelingsbehoeften, coaching en management development steunen bijvoorbeeld op het juist kunnen inschatten en beoordelen van de kwaliteiten van het aanwezige en toekomstige personeel. Indien goed toegepast, vormen assessment centers een interessant instrument om mensen te beoordelen en om een gedifferentieerd plaatje van de sterkten en zwakten van (toekomstige) medewerkers uit te tekenen. Het is dan ook niet verwonderlijk dat het laatste decennium de assessment-centermethode zich als een belangrijke techniek binnen HRM heeft geprofileerd.

In de voorgaande delen (kwaliteitsperspectief versus pragmatisch perspectief) zijn er duidelijke verschillen te ontwaren tussen hetgeen wetenschappelijk onderzoek over assessment centers suggereert en hetgeen er in de assessment-centerpraktijk gebeurt. We willen dit artikel dan ook afsluiten met een aantal aanbevelingen die de kwaliteit van het assessment center verbeteren en waarborgen. Daarbij is een belangrijk criterium dat de aanbevelingen implementeerbaar zijn. De praktijk moet er immers wel mee uit de voeten kunnen. We willen hierbij opmerken dat implementeerbaar vaak betekent dat het geld kost. Hoe meer van onderstaande aanbevelingen worden opgenomen, hoe meer het assessment center zal kosten. Deze kosten zitten verweven in alle fasen van het assessment-centerdesign. Met nadruk wijzen wij de lezer erop dat kwaliteit ook garant staat voor het nemen van kwalitatief betere beslissingen. Het tegenovergestelde is ook waar: een lage kwaliteit van het assessment center leidt tot het nemen van verkeerde beslissingen en die kosten vaak nog meer geld.

1) Inventarisatiefase

- Competenties zijn relevant voor de uit te voeren functietaken.
- Competenties hebben een eenduidige definitie.
- Competenties overlappen elkaar qua betekenis zo min mogelijk.
- Competenties zijn observeerbaar in gedrag.
- Competenties kunnen worden beoordeeld binnen het tijdsbestek van de opdracht.
- Het aantal te beoordelen competenties per opdracht is niet groter dan vijf.
- De (management)functie wordt afgedekt door de opdrachten.

- De opdrachten representeren de realiteit in voldoende mate en worden idealiter op maat gemaakt voor de organisatie.
- Deelnemers krijgen in de opdrachten voldoende kans om competentiegerelateerd gedrag te laten zien.

2) Operationalisatiefase

- Het assessorenteam bestaat uit zowel P&O- en lijnmanagers als psychologen.
- Assessoren zijn op de hoogte van beoordelaarseffecten zoals halo, strengheid of mildheid.
- Assessoren hebben voldoende ervaring met het beoordelen van assessment-centeropdrachten.
- Assessoren hebben inzicht in de vereisten van de functie en de cultuur van de organisatie.
- Assessoren zijn het eens over de betekenis van de competenties.
- Assessoren krijgen een frame-of-reference training.
- Rolspelers zijn in staat om op natuurlijke wijze vorm te geven aan de rol.
- Rolspelers zijn in staat om competentiegerelateerd gedrag uit te lokken door middel van het aanbieden van 'platforms'.

3) Uitvoeringsfase

- Assessoren roteren na iedere opdracht zodat iedere kandidaat slechts één keer wordt beoordeeld door iedere assessor.
- Vooral minder geroutineerde assessoren, zoals managers, hebben hulpmiddelen nodig bij het beoordelen, zoals observatiechecklists of gedragsverankerde beoordelingsschalen.
- Assessoren worden intensief en regelmatig getraind.
- Deelnemers (kandidaten) zijn bekend met de inhoud van het assessment center (inclusief de competenties zijn transparant).

4) Evaluatiefase

- Het assessment center wordt afgesloten met een assessorenvergadering, waar alle resultaten (simulaties, testen, vragenlijsten, interviews) gezamenlijk worden geëvalueerd.
- De validiteit en betrouwbaarheid van het assessment center worden frequent kwantitatief en kwalitatief geëvalueerd. De resultaten hiervan leiden tot verbeteringen in de assessment-centerprocedure.

Tabel 1. Checklist voor het waarborgen van de kwaliteit van het assessment center (Kolk, 2002)



Literatuur

- Ballantyne, I. en N. Povah, *Assessment and development centres*. Aldershot: Gower, 1995.
- Dam, K. van, W.M.M. Altink en B. Kok, 'De praktijk van de assessment center methode: Een inventarisatie van knelpunten'. *De Psycholoog*, 7, pp. 509-514, 1992.
- Fombrun, C., N.M. Tichy en M.A. Devanna, *Strategic Human Resource Management*. New York: Wiley, 1984.
- Kleinmann, M., C. Kuptsch en O. Köller, 'Transparency: A necessary requirement for the construct validity of assessment centres'. *Applied Psychology: An International Review*, 45, pp. 67-84, 1996.
- Kolk, N.J., *Assessment Centers: Understanding and improving construct-related validity*. Dissertatie: Vrije Universiteit, 2001.
- Kolk, N.J., 'Assessment center kan (nog) beter'. *Gids voor Personeelsmanagement*, 81 (5), pp. 32-37, 2002.
- Kolk, N.J., M.Ph. Born en H. van der Flier, 'The transparent assessment center: The effects of revealing dimensions to applicants'. *Applied Psychology; an International Review* (in druk).
- Kolk, N.J., M.Ph. Born en H. van der Flier, 'Impact of common rater variance on construct validity of assessment center dimension judgments'. *Human Performance* (in druk).
- Lievens, F., 'Assessor training strategies and their effects on accuracy, inter-rater reliability, and discriminant validity'. *Journal of Applied Psychology*, 86, pp. 255-264, 2001.
- Lievens, F., 'An examination of the accuracy of slogans related to assessment centres'. *Personnel Review*, 31, pp. 86-102, 2002.
- Lievens, F. en J.M. Conway, 'Dimension and exercise variance in assessment center scores: A large-scale evaluation of multi-trait-multimethod studies'. *Journal of Applied Psychology*, 86, pp. 1202-1222, 2001.
- Lievens, F. en H. Goemaere, 'A different look at assessment centers: Views of assessment center users.' *International Journal of Selection and Assessment*, 7, pp. 215-219, 1999.
- Lievens F. en H. Goemaere, 'Assessment centers in Vlaanderen: Onderzoek naar de kwantiteit en kwaliteit van toepassing'. *Tijdschrift voor HRM*, 3, pp. 53-68, 1999.
- Reilly, R.R., S. Henry en J.W. Smither, 'An examination of the effects of using behavior checklists on the construct validity of assessment center dimensions'. *Personnel Psychology*, 43, pp. 71-84, 1990.
- Sagie, A. en R. Magnezy, 'Assessor type, number of distinguishable categories, and assessment centre construct validity'. *Journal of Occupational and Organizational Psychology*, 70, pp. 103-108, 1997.
- Seegers, H.J.J.L., *Assessment Centers. Een personeelsinstrument voor de manager*. Deventer: Kluwer, 1997.
- Seegers, H.J.J.L. en J.L. Esser, *Talenten van vandaag voor organisaties van morgen: de Assessment Center Methode*. Deventer: Kluwer, 1982.
- Steltenpohl, E.L., 'Assessment centers - Management Development'. In: P.G.W. Jansen en F. de Jongh (red.), *Assessment centers: een open boek* (pp. 185-197). Utrecht: Het spectrum/marka, 1993.
- Thornton, G.C. III, *Assessment centers in Human Resource Management*. Reading, MA: Addison-Wesley, 1992.

$$\alpha^2 + \beta^2 = \gamma^2$$



BEOORDELEN IN COMPETENTIE-GERICHTE LEEROMGEVINGEN

MARIANNE ELSHOUT-MOHR, RON OOSTDAM EN MARIANNE OVERMAAT

In de onderwijsdoelstellingen die beroepsopleidingen zichzelf stellen, treedt voortdurend verandering op. Meestal gaat het hierbij om accentverschillen. Nu eens wordt het belang van gedegen vakkennis benadrukt, dan weer wordt een zwaarder accent gelegd op algemene vaardigheden, zoals het analyseren en oplossen van problemen en het mondeling en schriftelijk communiceren. Ook het idee dat studenten vooral moeten ‘leren leren’ krijgt af en toe ruime belangstelling (Peetsma en Volman, 1997). Een onderwerp dat op dit moment (weer) sterk in de aandacht staat, is de aansluiting tussen onderwijs en arbeid. Dit is ook het geval in het hoger beroepsonderwijs, waarop wij ons in deze bijdrage primair richten. Het gaat dan om de vraag hoe opleidingen studenten kunnen afleveren die over de competenties beschikken die de arbeidsmarkt van beginnende ‘professionals’ vraagt. Uitgangspunt van deze bijdrage is dat een (beroeps)opleiding er vanuit de bovengenoemde achtergrond voor heeft gekozen op een competentiegerichte wijze te gaan opleiden. Vragen die dan bovenkomen, en die we hier centraal stellen, zijn: (1) Hoe kan worden bewerkstelligd dat studenten niet alleen nuttige bouwstenen verwerven voor de beroepsuitoefening, maar ook de vaardigheid om deze te benutten in praktijksituaties en (2) wat zijn de consequenties voor de beoordeling van vorderingen en eindresultaten?

LEESWIJZER

Er zijn diverse mogelijkheden om tot competentiegericht opleiden over te gaan. Daarvan worden in deze bijdrage twee besproken, namelijk een instructivistische, die primair gericht is op de beroepsvaardigheden, en een constructivistische, die primair gericht is op de beroepsidentiteit. Ondanks de verschillen tussen deze beide opleidingsvarianten, hebben zij gemeenschappelijk dat een aanzienlijke wijziging vereist wordt in de wijze waarop studenten worden beoordeeld. Om de nodige ‘congruentie’ te behouden vereist competentiegericht opleiden een eigen beoordelingscultuur. Naar onze mening is het daarbij gewenst om een sterke samenhang aan te brengen tussen formatief en summatief beoordelen. Bij formatief beoordelen denken wij aan beoordelen in het kader van het leerproces, waarbij een goede balans dient te worden gevonden tussen beoordeling door docenten, medestudenten en zelfbeoordeling. Bij summatief beoordelen denken wij aan beoordeling van studenten ter selectie en certificering, bijvoorbeeld door assessorenpanels waarvan de leden niet alleen uit de opleiding maar ook uit de afnemende instellingen worden gerekruteerd. Voorts zijn wij van mening dat het leren omgaan met competentiestandaarden een integraal onderdeel van de beroepsopleiding zal moeten gaan vormen. Studenten dienen expliciet te worden voorbereid op een toekomst waarin van hen wordt ver-

wacht dat zij hun eigen competentieontwikkeling kritisch volgen, competenties van anderen kunnen beoordelen en een actieve bijdrage leveren aan het 'up to date' houden van de competentiestandaarden van het beroep. Op weg naar het inrichten van competentiegericht opleiden zijn er talloze valkuilen, waarvan sommige wellicht vermeden kunnen worden door een goede oriëntatie vooraf. Enkele valkuilen worden in deze bijdrage aan de orde gesteld.

TWEE VARIANTEN VAN COMPETENTIEGERICHT OPLEIDEN

In deze paragraaf bespreken wij twee varianten van competentiegericht opleiden. Het onderscheid tussen de twee varianten, waarin respectievelijk de beroepsvaardigheden en de beroepsidentiteit het centrale opleidingsdoel vormen, hebben wij gemaakt op basis van een omschrijving van het begrip competentie, die het ACOA (1997) heeft opgesteld. Eerst gaan we in op deze definitie, om vervolgens de twee varianten te schetsen.

Definitie van competentie

De ACOA-definitie luidt: *'Competenties: de vermogens van een individu waarmee opgaven van een beroep op een adequate, procesgerichte en productgerichte wijze kunnen worden aangepakt. Competenties zijn multidimensioneel, gestructureerd en onderling samenhangend.'* De ACOA noemt hier drie kenmerken van het competentiebep, die een verschillend gewicht kunnen krijgen bij het ontwerpen van competentiegericht opleiden.

In de eerste plaats wordt gesteld dat competenties *direct gerelateerd* zijn aan de opgaven van het beroep. Voorzover competenties berusten op kennis en (algemene) vaardigheden die tijdens een opleiding aan de orde komen, dient er dus een toespitsing op beroepspecifieke situaties plaats te vinden.

Ten tweede wordt aangegeven dat competenties meer omvatten dan het vermogen om adequaat te handelen. Het gaat uitdrukkelijk óók om het vermogen om *van*

voorkomen dat het competentiedenken dat zij voorstaan, wordt vereenzelvigd met een 'behavioristische', puur handelingsgerichte, benadering van beroepscompetenties (zie voor behandeling van dit onderwerp O'Neil, 1997; Verloop, 1995). Voor opleidingen betekent dit kenmerk dat men het belang van kennis en vaardigheden (wij spreken verder van 'expertkennis') niet dient te onderschatten, maar dat men wel een duidelijke, aan de competentieontwikkeling gerelateerde plaats moet geven aan studieonderdelen waarin deze expertkennis centraal staat. In het oog moet worden gehouden dat expertkennis niet alleen moet worden toegepast, maar ook functioneel moet zijn bij reflectie op eigen functioneren en in de communicatie met anderen.

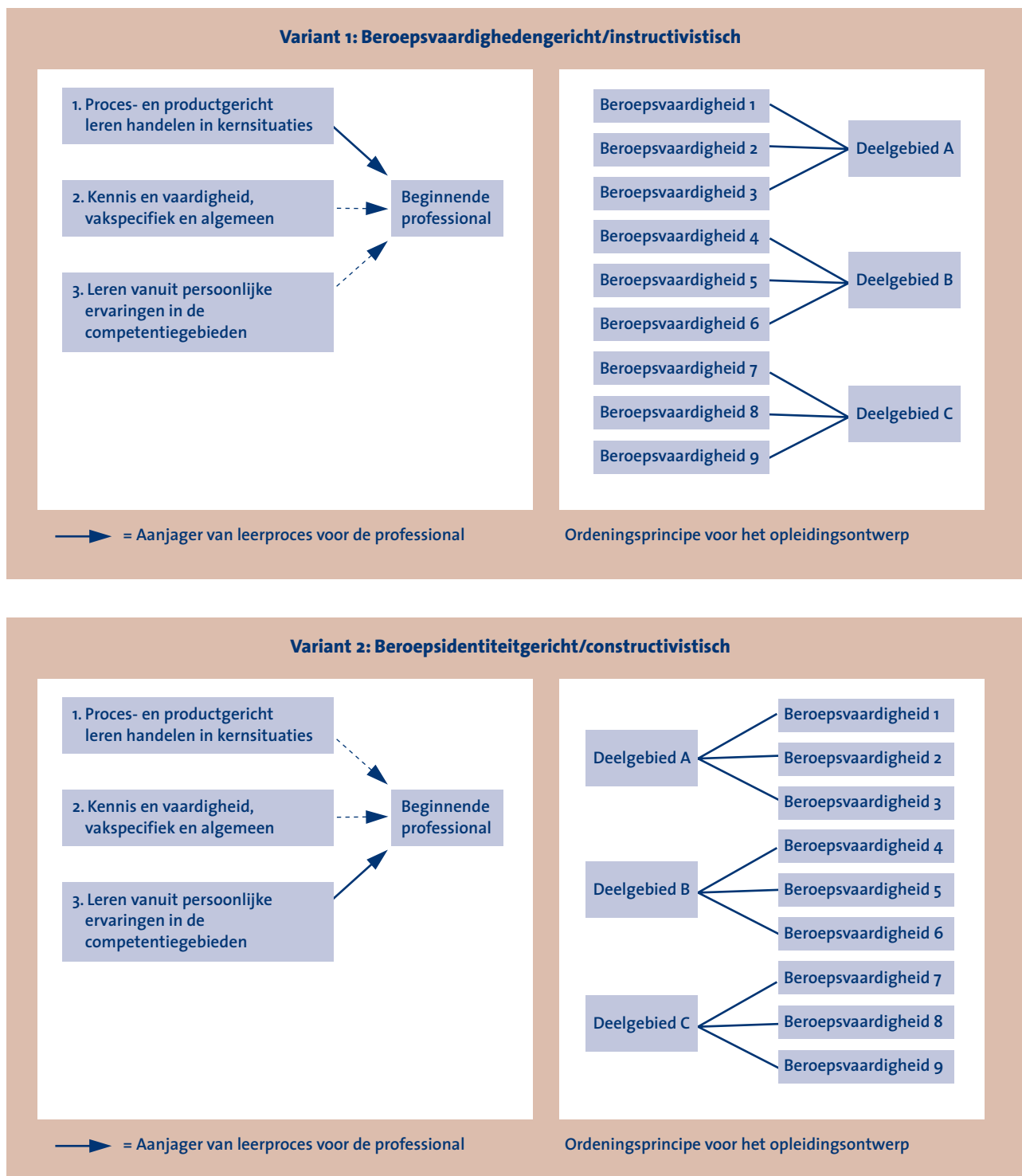
Een derde punt dat in de ACOA-omschrijving van het competentiebep, afzonderlijk wordt benadrukt, is dat de *verschillende competenties deel uitmaken van een multidimensioneel geheel*. Zij vormen met elkaar een complex netwerk van kennis, vaardigheden, normen, waarden, voorkeuren en rolopvattingen waarop de persoon het professionele handelen baseert. Dit multidimensionele netwerk is voor elke beroepsbeoefenaar verschillend, en bepaalt hoe de opgaven van het beroep worden benaderd vanaf het moment van waarnemen en interpreteren van de situatie tot en met het moment van uitvoeren, evalueren en reflecteren. Dit multidimensionele netwerk wordt ook wel 'competency' genoemd ter onderscheiding van 'competence(s)'. Mulder (2002) spreekt ook van de algemene bekwaamheid, ter onderscheiding van de competenties die de specifieke bekwaamheden aanduiden. Wij hebben gekozen voor de term 'beroepsidentiteit', omdat daarin niet alleen de complexe structuur van de 'competency' tot uitdrukking komt, maar ook de multidimensionaliteit. Opleidingen zullen er oog voor dienen te hebben dat de ervaring van de student op diverse competentiedeelgebieden toeneemt in de loop van de opleiding. Er treedt versterking en consolidatie op van sommige elementen in het multidimensionele netwerk en bijstelling en transformatie van andere elementen. Tot de beroepscompetentie wordt door velen ook de 'doorgroeicompetentie' gerekend. Dit betekent dat mensen ook in latere fasen van de beroepsuitoefening continu aan de identiteitsontwikkeling blijven werken (zie Korthagen et al., 2002; Melief et al., 2002).

Opleidingsvarianten

In figuur 1 worden *twee opleidingsvarianten* schematisch weergegeven waarin verschillende accenten worden gelegd op de drie kenmerken. Bij de eerste variant, *beroepsvaardigheden gericht/instructivistisch*, ligt het accent op de kenmerken 'directe koppeling aan de opgaven van het beroep' en 'het vermogen vanuit de nodige afstand het eigen handelen te beschouwen en

De verschillende competenties vormen een complex netwerk van kennis, vaardigheden, normen, waarden, voorkeuren en rolopvattingen

uit de nodige afstand het eigen handelen te beschouwen en bespreekbaar te maken. Dit blijkt uit de zinsnede 'adequate, procesgerichte en productgerichte wijze', die vermoedelijk door de ACOA is toegevoegd om te



Figuur 1. Twee varianten van competentiegericht opleiden: beroepsvaardighedengericht (instructivistisch) versus beroepsidentiteitgericht (constructivistisch)

bespreekbaar te maken'. Bij de tweede variant, *beroepsidentiteit gericht/constructivistisch*, ligt het accent op het gegeven dat competenties deel uitmaken van een multidimensioneel geheel.

De term '*beroepsvaardigheden*' in figuur 1 verwijst naar de vaardigheden die nodig zijn om adequaat te func-

tioneren in kenmerkende en kritische beroepssituaties. Onderdeel van dit adequaat functioneren is dat de persoon kan uitleggen en verantwoorden wat hij doet en waarom. Vertaald in opleidingstermen kan gesteld worden dat de directe koppeling van competenties aan de beroepspraktijk impliceert dat stages en (eventueel gesimuleerde) beroepssituaties belangrijke

instructievormen zijn. Deze dienen zodanig te worden geïntegreerd in de opleiding dat er geen breuk ontstaat tussen het formele ‘weten wat’ en ‘weten hoe’ en de praktische kennis die studenten ontwikkelen tijdens het werken in de beroepspraktijk. Dat studenten het beroepsmatig handelen moeten kunnen expliciteren en verantwoorden, impliceert dat er in de begeleiding en studentinteracties veel ruimte dient te zijn voor verdiepende communicatie over praktijkervaringen.

De term ‘*beroepsidentiteit*’ in figuur 1 verwijst naar het multidimensionele karakter van het complexe netwerk dat aan de beroepscompetentie ten grondslag ligt. Dit dient vorm te krijgen in de opleiding door studenten veelvuldig in beroepssituaties te brengen waarin naast kennis en vaardigheden ook hun normen en waarden, voorkeuren en rolopvattingen in het geding zijn. Door vervolgens het functioneren in deze situaties tot onderwerp van discussie en reflectie te maken, wordt de ontwikkeling van de beroepsidentiteit een centraal onderdeel van de studie. Tevens nodigt dit studenten uit om in te gaan op de relatie tussen de eigen beroepsidentiteitsontwikkeling en datgene wat aan de orde komt in de aangeboden informatiebronnen (zoals colleges en studieboeken) en trainingen (zoals vaardigheidslessen en practica).

In het *linkerdeel van figuur 1* is te zien dat in beide varianten dezelfde componenten een rol spelen en dat het doel in beide gevallen is dat de student de opleiding verlaat als beginnend professional. De componenten zijn weergegeven in blok 1, 2 en 3. Blok 1 verwijst naar de noodzaak om studenten gelegenheid te bieden voor proces- en productgericht handelen in kernsituaties van het beroep. Blok 2 geeft aan dat er studieonderdelen dienen te zijn waarin de kennis en vaardigheden aan de orde komen waarover beginnen de professionals dienen te beschikken. Blok 3 laat zien dat studenten de gelegenheid dienen te krijgen om vanuit de eigen ervaringen met het functioneren in de competentiegebieden aan de algemene beroepsbekwaamheid te werken.

In het *rechterdeel van figuur 1* is de beoogde startbekwaamheid van de beginnende professional ontleed in beroepsvaardigheden (aangeduid als kleine ovalen) en deelgebieden waarop de beroepsidentiteit van de studenten vorm (moet) krijgen (de grote ovalen). Binnen een lerarenopleiding kunnen bijvoorbeeld als beroepsvaardigheden of ‘kerncompetenties’ worden genoemd ‘de leraar draagt zorg voor een veilige leeromgeving voor leerlingen’, ‘de leraar draagt bij aan realisering en ontwikkeling van het pedagogisch klimaat in team/organisatie’, et cetera. Bij deelgebieden kan gedacht worden aan ‘pedagogisch’, ‘vakdidactisch’, ‘organisatorisch’ en ‘interpersoonlijk’ handelen.

De volgorde waarin de beroepsvaardigheden en deelgebieden zijn gepresenteerd, is in beide varianten verschillend, en ook is er in beide varianten een andere zwarte, doorgetrokken pijl geschetst.

De zwart doorgetrokken *pijl in variant 1* geeft aan dat het proces- en productgericht handelen (blok 1) de aanjager van het leerproces vormt. De positionering en verbindingslijnen tussen de beroepsvaardigheden en de deelgebieden bij deze variant geven aan dat de beroepsvaardigheden het primaire doel zijn van de opleiding en dat de identiteitsontwikkeling op de deelgebieden wordt gezien als een logisch gevolg op het leren beheersen van deze vaardigheden. Wij noemen de variant *instructivistisch*, omdat de opleiders het voortouw nemen bij het vaststellen van de studie-inhouden. Welke kerncompetenties aan de orde komen en aan welke standaarden studenten zullen moeten voldoen, wordt al bepaald voordat studenten aan het leerproces beginnen. Uiteraard betekent deze docentsturing niet dat actieve deelname van studenten overbodig zou zijn. De zwart doorgetrokken *pijl in variant 2* geeft aan dat de eigen ervaringen die studenten opdoen tijdens het functioneren in de deelgebieden (blok 3) de aanjager van het leerproces vormen. De positionering en de verbindingslijnen tussen de deelgebieden en de beroepsvaardigheden geven aan dat de identiteitsontwikkeling op de deelgebieden het primaire doel vormt van de opleiding en dat de ontwikkeling van de beroepsvaardigheden wordt gezien als een logisch gevolg. Deze variant typeren wij als *constructivistisch*, omdat elke student uitgaat van de eigen ervaringen en het eigen leerproces dient te sturen. Uiteraard is een actieve begeleiding van docenten hierbij onmisbaar. In de praktijk zal dat betekenen dat ook in deze variant de docenten het voortouw nemen. Zij zullen hun inspanningen echter richten op de planning van de wijze waarop studenten ervaringen gaan opdoen in praktijksituaties (gevarieerd naar deelgebieden) en op begeleiding van keuze- en reflectieprocessen van studenten. Ook de voorbereiding van het inhoudelijke onderwijsaanbod zal bij de constructivistische variant anders zijn dan bij de instructivistische tegenhanger. Het aanbod zal er sterker op zijn afgestemd dat studenten individuele leerroutes volgen en op verschillende momenten met eigen opleidingsvragen zullen komen (Elshout-Mohr en Oostdam, 2001; Elshout-Mohr, Oostdam en Overmaat, 2001).

HET BEOORDELEN VAN STUDENTEN IN EEN COMPETENTIEGERICHTE OPLEIDING

In een competentiegerichte leeromgeving zullen studenten anders moeten worden beoordeeld: via andere methoden, waarbij congruentie met de opleidingsvariant van belang is.

Beoordelingsmethoden

Verschillen tussen *opleidingsvarianten* zullen niet alleen doorwerken in het onderwijsaanbod en de begeleiding, maar ook in de *wijze waarop studenten worden beoordeeld*. De invoering van competentiegericht onderwijs vraagt om nieuwe beoordelingsinstrumenten en -procedures. Nagegaan moet immers worden of de ontwikkeling van de beroepsidentiteit en de beroepsvaardigheden van studenten aan bepaalde maatstaven voldoen en of verworven kennis en vaardigheden door studenten op de beoogde, competentiegerelateerde wijze worden ingezet.

Bij een keuze voor *variant 1* ligt het voor de hand om allereerst de aandacht te richten op het selecteren en benoemen van beroepsvaardigheden (kerncompetenties) en het ontwikkelen van methoden om de beheersing ervan te meten. Men kan de vraag nog even voor zich uitschuiven hoe men uiteindelijk waar gaat maken dat de opleiding ook werkelijk competentiegericht is, in die zin dat de studenten aan het einde van de rit als ‘beginnend professional’ kunnen functioneren en al enigermate een beroepsidentiteit hebben ontwikkeld. Ook de vraag hoe men traditionele toetsen op beroepsspecifieke kennis en vaardigheden zou kunnen herzien om ze optimaal te laten passen bij een competentiegericht opleidingsconcept lijkt bij een keuze voor variant 1 niet direct urgent.

Bij een keuze voor *variant 2* dient zich in eerste instantie vooral de vraag aan hoe (met welke instrumenten en methoden) de ontwikkeling van een beroepsidentiteit van studenten kan worden beoordeeld. Bij deze benadering komen vragen over het beoordelen van afzonderlijke kerncompetenties op het tweede plan en lijkt het, net als bij de eerste variant, misschien niet direct noodzakelijk om veranderingen aan te brengen in traditionele toetsen op beroepsspecifieke kennis en vaardigheden. Hoewel het natuurlijk prettig kan zijn om niet te veel problemen tegelijkertijd aan te pakken, is een goede oriëntatie zeker gewenst op wat er in latere stadia van de innovatie nog moet gebeuren. Ook beoordelingsproblemen die niet direct aan de orde lijken te zijn, dienen te worden meegewogen in de beslissingen en de innovatiestrategie.

Congruentie

Met de keuze voor een competentiegericht opleidingsmodel gaan opleidingsinstellingen een traject in waarin zij uiteindelijk, ook als beoordelaars, met alle kenmerken van het competentiebegrip te maken krijgen. Het zou daarom verstandig kunnen zijn om van meet af aan te sturen op een beoordelingscultuur die congruent is met alle kenmerken van competenties, namelijk 1) dat competenties direct aansluiten op beroepssituaties, 2) dat studenten het eigen handelen moeten kunnen expliciteren en verantwoorden (met

behulp van een goede kennis- en vaardighedenbasis) en 3) dat competenties deel uitmaken van de beroepsidentiteit, een multidimensioneel netwerk van kennis, vaardigheden, normen, waarden, voorkeuren en rol-opvattingen. Enkele grote lijnen van een dergelijke beoordelingscultuur zijn rechtstreeks uit deze kenmerken af te leiden.

Het eerste kenmerk (*competenties sluiten direct aan op beroepssituaties*) wijst bijvoorbeeld duidelijk in de rich-

Het leren beoordelen van zichzelf en anderen zal een belangrijkere plaats krijgen in opleidingen

ting van een relatief open beoordelingscultuur, waarin het oordeel over het functioneren van studenten niet exclusief de taak is van opleiders. Aangezien diverse personen, elk vanuit een eigen rol, informatie hebben over de adequaatheid van het handelen van studenten in beroepssituaties en aangezien zij de passendheid en effectiviteit van het handelen ook vanuit hun eigen perspectief zullen waarderen, ligt het voor de hand om diverse personen een rol te geven in de beoordeling. Ten behoeve van de formatieve beoordeling kunnen beoordelaars elkaar op natuurlijke wijze afwisselen en aanvullen. Op summatieve beoordelingsmomenten kan waarschijnlijk het best worden gewerkt met beoordelingspanels, waarin verschillende relevante perspectieven systematisch vertegenwoordigd zijn.

Ook het tweede kenmerk (*competente personen kunnen het handelen expliciteren en verantwoorden*) leidt tot de conclusie dat de beoordelingscultuur wellicht een meer open karakter zal krijgen. Het is waarschijnlijk onvoldoende dat opleiders expertkennis van studenten toetsen in afzonderlijke tentamens en voor toetsdoeleinden geconstrueerde opdrachten. Aanvullend zal men willen beoordelen of studenten relevante kennis en vaardigheden ook daadwerkelijk inzetten bij het beroepsmatig handelen en of zij er gebruik van weten te maken bij het expliciteren en verantwoorden van hun handelen ten overstaan van bijvoorbeeld medestudenten of andere betrokkenen.

Het derde kenmerk (*competenties maken deel uit van de beroepsidentiteit*) versterkt de algemene indruk dat ingrijpende veranderingen in de beoordelingscultuur nodig zijn als instellingen het werkelijk serieus nemen met de competentiegerichtheid. Recht doen aan individuele verschillen in beroepsidentiteit vraagt om vervanging van ‘standaardgestuurd beoordelen’ door ‘assessment’ (Elshout-Mohr en Oostdam, 2001). Terwijl bij ‘standaardgestuurd beoordelen’ het accent ligt op het meten van prestaties met behulp van vaste maatstaven, van tevoren geformuleerde (en bekendgemaakte)

eisen en voor alle kandidaten gelijke gedragsindicatoren, worden de maatstaven die bij een 'assessment' worden aangelegd mede bepaald door de inbreng van de kandidaat. De assessmentsituatie is vergelijkbaar met een sollicitatie: verschillende personen kunnen om verschillende redenen geschikt zijn voor een functie. In een beoordelingscultuur waarin assessments worden afgenomen is het passend dat de student eerst zelf informatie verschaft over de eigen beroepsidentiteit, die uiteraard nog in ontwikkeling is. Ook geeft de student in een zelfevaluatie aan waarom hij meent aan relevante standaarden te voldoen. Vervolgens selecteert en formuleert een assessmentpanel vragen en opdrachten die hen in staat zullen stellen de competenties van de student vast te stellen en om daarover terugkoppeling te geven (formatief). In een eventuele beslissing over de student (summatief) wordt rekening gehouden met zowel de beroepsidentiteit als de antwoorden op gestelde vragen en de prestaties bij opdrachten.

We concluderen dat in de beoordeling, ter wille van het bereiken van congruentie tussen (één van beide opleidingsvarianten voor) competentiegericht opleiden en beoordelingsmethoden, allerlei nieuwe elementen moeten worden aangebracht waarover nog relatief weinig bekend is. Enkele instrumenten die wel al voorhanden zijn, worden kort besproken in de volgende paragraaf. Wij concluderen voorts dat instellingen een echte omslag in de beoordelingscultuur misschien wel kunnen uitstellen, maar dat zij er uiteindelijk waarschijnlijk toch niet aan ontkomen. Bij een voorkeur voor variant 1 is het belangrijk dat instellingen zich realiseren dat zij uiteindelijk niet zullen kunnen volstaan met een beoordelingssysteem dat voorbijgaat aan het aspect van beroepsidentiteit, dat (naar onze mening) onlosmakelijk aan het competentiebegrip is verbonden. Evenzeer geldt dat een keuze voor variant 2 instellingen op de wat langere termijn niet zal ontslaan van de noodzaak om het niveau van bepaalde afzonderlijke beroepsvaardigheden van studenten vast te stellen. In beide varianten, tenslotte, zullen ook voor het beoordelen van de expertkennis passende vormen moeten worden gevonden (Elshout-Mohr, Oostdam en Overmaat, in druk; Overmaat et al., 2001).

Het bovenstaande geeft een eerste indruk van het soort beoordelingscultuur dat zou passen bij competentiegericht opleiden. Wij verwachten een toenevende rol van beoordelaars van buiten de opleiding en eveneens een toenemende rol van zelfbeoordeling door studenten en 'peer'-beoordeling door medestudenten, zowel in formatieve (ontwikkelingsgerichte) als in summatieve (beslissingsgerichte) beoordelings-

situaties (zie ook Snoek, 2000). Voorts verwachten wij dat het leren beoordelen van zichzelf en anderen een belangrijker plaats in zal gaan nemen in de opleidingen. Pogingen om studenten sterker in beoordelingsprocedures (van henzelf of medestudenten) te betrekken, kunnen slechts slagen als studenten de gelegenheid krijgen om systematisch en kritisch te leren omgaan met beoordelingsmaatstaven en criteria (zie ook Dochy, Segers en Sluijsmans, 1999).

BEOORDELINGSINSTRUMENTEN EN -PROCEDURES

Er zijn in de laatste decennia verscheidene nieuwe beoordelingsinstrumenten ontworpen en in gebruik genomen, waaronder het *portfolio*, de *kritische situatie* en de *voortgangstoets*. Figuur 2 geeft een overzicht van de wijze waarop deze instrumenten deel uit kunnen maken van een voorstelbare en wellicht effectieve beoordelingscultuur. Veel van de ideeën in figuur 2 zijn geïnspireerd door de werkwijze die ontwikkeld is door de Educatieve Faculteit Amsterdam (EFA), die in 1997 de weg is ingeslagen in de richting van competentiegericht opleiden. We belichten de drie genoemde methoden, en schetsen vervolgens hoe naar ons idee beoordelen steeds meer een onderdeel wordt van het leerproces, waardoor doorlopende leerlijnen ontstaan.

Portfolio

Onder een *portfolio* verstaan we het product dat ontstaat als een persoon uit een zelf gevulde, al dan niet elektronische, vergaarbak van allerlei plannen, producten (werkstukken, verslagen, video-opnamen van eigen handelen in beroepssituaties, en dergelijke) en ideeën, naar wens en behoefte een selectie maakt, daarin structuur aanbrengt en het geheel presentabel maakt ten behoeve van een doel (Elshout-Mohr, 2002). In Figuur 2 worden diverse functies van het portfolio genoemd. De kandidaat kan bijvoorbeeld eigen competenties presenteren, in de vorm van een zelfevaluatie die wordt ondersteund door evaluaties van anderen. De kandidaat kan het portfolio ook zo inrichten dat het een basis vormt voor een presentatie van de eigen beroepsidentiteit. Daartoe zullen bepaalde producten een goed toegankelijke plaats in het portfolio moeten krijgen en kan het ook zinvol zijn om toekomstplannen in directe aansluiting op deze producten te presenteren, zodat de assessoren snel een goed beeld kunnen krijgen van een doorgaande leerlijn. Als het doel tevens is om assessoren in de gelegenheid te stellen zich een oordeel te vormen over de producten, vereist dit van de kandidaat eveneens de nodige voorbereiding. Producten moeten worden voorzien van 'kopjes' waarin kort wordt samengevat wat de aard en herkomst is, in welke context het product is gemaakt

	Beoordelen als onderdeel van de instructie	Beoordelen met een formatief doel (ontwikkelingsgericht)	Beoordelen met een summatief doel (beslissingsgericht)
Beroepsvaardigheden (kerncompetenties)	Studenten schrijven zelfevaluaties naar aanleiding van hun handelen in beroepssituaties en geven elkaar daarop feedback.	Studenten verzamelen evaluaties van diverse personen die bij hun handelen in beroepssituaties zijn betrokken, waaronder stagebegeleiders uit de opleiding. Studenten gaan na welke competenties zij hebben versterkt en schrijven zelf-evaluaties waarop docenten feedback geven.	Assessoren baseren hun oordeel op: 1. beschrijvingen die studenten in het <i>portfolio</i> geven van hun competenties, op basis van zelfevaluatie en met toevoeging van bewijzen in de vorm van evaluaties van anderen; 2. de prestaties van studenten in <i>kritische situaties</i> (met name tijdens het uitvoeringsdeel van deze situaties).
‘Expertkennis’ die nodig is voor expliciteren en verantwoorden van handelen in beroepssituaties	Studenten beoordelen, naar aanleiding van bestudering van expertkennis (het formele ‘know-how’ en ‘know-what’), de relatie met de eigen ervaringen in de beroepspraktijk en de mogelijke waarde van de kennis voor het expliciteren en verantwoorden van (eigen) handelen in beroepssituaties. Studenten werken voorbeelden uit van goede explicitering en verantwoording van (eigen) handelen en leggen die aan elkaar ter beoordeling voor.	Docenten sturen het proces (zie vorige kolom) bij door actieve deelname aan plenaire discussies waarin studenten de bevindingen samenvatten en uitgewerkte voorbeelden presenteren. Studenten maken periodiek diagnostische voortgangstoetsen om zicht te krijgen op de eigen beheersing van expertkennis. Op basis van het resultaat stuurt de student zichzelf bij, eventueel met behulp van advies van docenten.	Assessoren baseren hun oordeel op: 1. de behaalde scores bij de laatste <i>voortgangstoets</i>, die zijn opgenomen in het <i>portfolio</i>; 2. de prestaties van de studenten in <i>kritische situaties</i> (en met name tijdens het plannings- en reflectiedeel van deze situaties).
Beroepsidentiteit: multidimensioneel netwerk van kennis, vaardigheden, normen, waarden, voorkeuren en rolopvattingen	Studenten raadplegen elkaar en portfolio’s van ouderejaars om ideeën op te doen omtrent de wijze waarop de beroepsidentiteit zichtbaar kan worden gemaakt en omtrent de criteria waaraan een portfolio moet voldoen om een overtuigend (authentiek, betrouwbaar, up-to-date) beeld te kunnen geven van de beroepsidentiteit en ontwikkelingen daarin.	Studenten werken continu aan een (elektronisch) portfolio, waarin zij materiaal verzamelen aan de hand waarvan zij zichzelf kunnen presenteren als beginnend professional. Op de opzet en inhoud van dit portfolio krijgen zij regelmatig feedback van docenten, medestudenten en anderen die de studenten daartoe zelf kunnen uitnodigen.	Assessoren baseren hun oordeel op: 1. een <i>portfoliopresentatie</i> waarin de student uiteenzet wat zijn/haar portfolio bevat t.a.v. de eigen beroepsidentiteit; 2. hun eigen evaluatie van het <i>materiaal</i> in het portfolio; 3. een beoordeling van de kwaliteit van een ‘<i>interviewgesprek</i>’ over het functioneren van de student in een <i>kritische situatie</i> (tijdens alle drie delen van deze situaties).

Figuur 2. Componenten van een beoordelingscultuur in een competentiegerichte opleiding

en waar de beschouwer in het bijzonder op zou kunnen letten. Bij een tekort aan ordening en structuurelementen is een portfolio voor assessoren ontoegankelijk en kan het ook nauwelijks een functie vervullen in de begeleiding of beoordeling (Wielenga, 2002).

Kritische situatie

Onder een *kritische situatie* verstaan wij een situatie die een beroep doet op een kerncompetentie van de kandidaat. De situatie, bestaande uit een situatiebeschrijving en een opdracht, wordt ten tijde van het assessment aan de kandidaat voorgelegd (meestal in de vorm van een simulatie). De kandidaat krijgt een afgepaste hoeveelheid tijd (bijvoorbeeld twee dagdelen) om aan de opdracht te werken en van dit werk een neerslag (schriftelijk, eventueel aangevuld met audio- of videoband) te maken. In figuur 2 wordt (in kolom 3) verwezen naar drie delen van de kritische situatie, namelijk een planfase, een uitvoeringsfase en een reflectiefase. Deze indeling is ontleend aan Sandberg et al. (1999). In de planfase kan de kandidaat tonen of hij beschikt over de kennis en vaardigheid om de situatie te analyseren en aan te geven wat hij zou doen, in welke volgorde en waarom. In de uitvoeringsfase kan de kandidaat tonen of hij de geplande (en eventuele andere adequate) handelingen kan uitvoeren in een praktijk situatie. In de reflectiefase kan de kandidaat datgene wat hij in de praktijkfase heeft gedaan, relateren aan meer algemene principes die ook in andere praktijk situaties van toepassing zijn (transfer). In de laatste kolom van figuur 2 is aangegeven dat de kritische situatie voor verschillende summatieve beoordelingsdoelen kan worden ingezet, waarbij de assessoren speciaal letten op de prestaties van de kandidaat in een bepaalde fase van werken aan de kritische situatie. Een intervisiegesprek kan worden opgevat als een speciale toepassing van de kritische situatie. In zo'n gesprek wordt niet zozeer gelet op de kwaliteit van het handelen in de afzonderlijke fasen, maar op de wijze waarop de kandidaat over het handelen communiceert. Is de kandidaat in staat sterke en zwakke punten in het totaal (planning, uitvoering en reflectie) aan te wijzen en te benoemen en is de kandidaat in staat nieuwe aanknopingspunten of thema's aan te wijzen voor verdere ontwikkeling van de beroepsidentiteit?

Voortgangstoets

Een derde instrument dat een bijdrage kan vormen tot competentiegericht beoordelen is de *voortgangstoets* ter meting van de expertkennis (Driessen en Van der Vleuten, 2000). Onder een voortgangstoets verstaan we een toets die bestaat uit een grote verzameling traditionele toetsvragen en -opgaven, die met elkaar de expertkennis dekken die de studenten in de loop van

de studie dienen te verwerven. Een dergelijke toets is een voorbeeld van een toets die voortkomt uit een instructivistische onderwijs traditie, maar toch goed gebruikt kan worden in een constructivistische leeromgeving (Elshout-Mohr, Oostdam en Overmaat, in druk). Studenten die verschillende leer routes volgen en een verschillende beroepsidentiteit ontwikkelen, kunnen allen dezelfde toets gebruiken om vast te stellen of zij voldoende expertkennis verwerven en dit ook kunnen tonen.

Doorlopende leerlijnen

De openvolging van de kolommen in figuur 2 laat schematisch zien dat er doorlopende lijnen kunnen zijn tussen beoordelen als onderdeel van het instructieproces (ten dienste van het leren beoordelen), beoordelen in formatieve zin en beoordelen in summatieve zin. In het gunstigste geval sluiten summatieve beoordelingen door assessorenpanels (zoals aangegeven in kolom 3) goed aan bij formatieve beoordelingen door docenten, stagebegeleiders en medestudenten (kolom 2). In een efficiënte beoordelingscultuur sluiten beide beoordelingstypen (summatief en formatief) bovendien direct aan op activiteiten die studenten en hun docenten en begeleiders al hebben ondernomen om het leerproces vorm te geven (kolom 1). Hoewel figuur 2 de indruk zou kunnen wekken dat het beoordelen in een competentiegericht opleiding heel ingewikkeld is en veel tijd en inspanning kost, is het voorbarig om op die eerste indruk af te gaan. Er vallen immers ook traditionele beoordelingsmomenten weg, zoals de puur summatief en controlerend bedoelde afsluitende toetsen aan het einde van modules. Bovendien komen alle activiteiten uit figuur 2, mede door hun onderlinge samenhang, tamelijk direct ten goede aan het proces van competentieverwerving dat het doel vormt van de opleiding. Tijd die besteed wordt aan beoordelen is in dat geval goed bestede tijd (Black en William, 2001).

HET TRAJECT IN DE RICHTING VAN COMPETENTIEGERICHT OPLEIDEN EN BEOORDELEN

Het inrichten van een competentiegericht opleiding met bijpassende beoordelingsprocedures is op zichzelf al een complexe taak, maar deze wordt in de praktijk nog bemoeilijkt doordat men niet vanuit het niets begint. Er is al een opleiding, er zijn al begeleidingsvormen en toetsen en de verschillende leden van de staf hebben hiermee een bepaalde binding, zowel emotioneel als zakelijk. Bovendien geldt dat men bestaande 'congruenties' doorbreekt als men – stap voor stap – langs de weg der geleidelijkheid veranderingen probeert door te voeren. Het congruentieprin-

cipe houdt echter niet alleen in dat onderwijs het meest effectief verloopt wanneer alle componenten van dat onderwijs zijn 'geënt' op eenzelfde leertheorie, zoals een 'instructivistische' of een 'constructivistische' leertheorie (Biggs, 1999). Het houdt ook in dat doorbreken van de congruentie aanleiding geeft tot ergernis, onbegrip en weerstand. We presenteren vanuit het congruentieperspectief enkele overwegingen voor en tegen beide varianten

Voor instellingen die een traditionele (instructivistische) opleiding willen omvormen tot een competentiegericht, lijkt het in eerste instantie het meest eenvoudig en haalbaar om te kiezen voor een instructivistische variant, waarin primair wordt gewerkt aan de beroepsvaardigheden van studenten. Wil men in het afsluitende assessment van de studenten echter ook de beroepsidentiteit van de studenten betrekken, dan ontstaan er problemen. Op dat moment vereist hetzelfde congruentieprincipe (dat men aanvankelijk meende te kunnen dienen) namelijk dat de 'nieuwe' beoordelingsvormen (waar op dat moment onvermijdelijk naar gegrepen moet worden) niet uit de lucht komen vallen. Een portfolio bijvoorbeeld is een waardevol beoordelingsinstrument voor studenten die er gedurende voorgaande jaren mee hebben leren werken, maar het is niet goed hanteerbaar voor studenten die langdurig gefunctioneerd hebben in een instructivistische, docentgestuurde leeromgeving.

Instellingen die een overstap naar competentiegerichtheid plannen met een goed oog voor de lange termijn, zullen daarom wellicht (zij het met de nodige aarzeling) de voorkeur geven aan een constructivistische variant, zoals weergegeven in het tweede deel van figuur 1. In dat geval wordt men weliswaar op een directer wijze geconfronteerd met een cultuuromslag, maar de opleiding kan wel van meet af aan congruent gemaakt worden met het portfolio als beoordelingsinstrument in de afsluitende fase van de studie.

In de regel lijkt een keuze voor een opleidingsconcept dat gebaseerd is op een constructivistische theorie over leren ons, vroeger of later, onvermijdelijk voor hbo-opleidingen die zich willen presenteren als volledig competentiegericht. Binnen een opleiding die constructivistisch is ingericht kunnen overigens wel episodes worden ingebouwd die op een instructivistische theorie zijn gebaseerd. Denk bijvoorbeeld aan episodes waarin studenten zich volgens een door docenten ontworpen opdrachtencyclus een kernvaardigheid eigen maken, zoals het verzorgen van een mondelinge presentatie. Dergelijke episodes kunnen worden afgesloten met bijpassende (instructivistische) toetsen, die eruit bestaan dat de student de verworven vaardigheid demonstreert en daarbij probeert zo goed moge-

lijk te voldoen aan de standaarden die de docent tijdens de oefenfase onder de aandacht heeft gebracht. Wil men het constructivistische karakter van de opleiding niet te zeer aantasten, dan dienen dergelijke ingebouwde episodes echter wel beperkt in omvang en aantal te blijven en liefst facultatief te worden aangeboden. In het ideale geval kan elke student voor zichzelf bepalen of en wanneer een bepaalde episode in de persoonlijke leerroute past.

VALKUILEN BIJ HET MAKEN VAN DE OVERSTAP NAAR COMPETENTIEGERICHT BEOORDELEN

Het overstappen op competentiegericht beoordelen heeft slechts kans van slagen als alle betrokkenen voorbereid zijn op een proces van lange adem. Er zullen onvermijdelijk enkele cycli moeten worden doorlopen om de problemen die zich voordoen te identificeren en er oplossingen voor te vinden. We belichten drie van deze veelvoorkomende problemen of valkuilen.

Voldoende middelen voor de startfase, niet voor daarna

De grootste valkuil is dan ook dat men wel voldoende innovatiemiddelen (geld, menskracht) weet te verwerven voor een behoorlijke startfase, maar dat men te weinig reserveert voor de fase van het verfijnen van procedures en het oplossen van problemen. Dit heeft dan tot gevolg dat men op een gegeven moment moet gaan kiezen voor compromissen en ad hoc oplossingen die ertoe leiden dat de innovatie haar elan verliest. Men blijft steken in uitwerkingen die hybride zijn (een mix van het oude en het nieuwe). De grote lijn van de vernieuwing dreigt uit zicht te verdwijnen en naast innovatiemoeheid kan ook nostalgisch verlangen ontstaan naar vroegere tijden.

Een concrete vorm waarin bovengenoemde valkuil tot uiting komt is onder andere waarneembaar binnen de EFA. Hier was gekozen voor een innovatievorm die de verschillende opleidingen (bijvoorbeeld de opleidingen tot leraar in de talen of de wiskunde) de gelegenheid gaf zelf te experimenteren met eigen toetsvormen in de eerste jaren waarin het competentiegericht opleidingsmodel werd ingevoerd. Na deze periode zouden de ervaringen worden gebundeld om vervolgens te komen tot een vorm die voor alle opleidingen zou gaan gelden. Dit geplande vervolg op de experimentele fase is echter niet goed van de grond gekomen, hetgeen voor een belangrijk deel moet worden toegeschreven aan gebrek aan geld. Het lukte niet hiervoor de nodige tijd vrij te maken. Ook de experimenteerfase als zodanig had overigens al te lijden aan tijdgebrek. Zo is het bijvoorbeeld niet in alle EFA-oplei-

dingen gelukt een goede vorm te vinden voor het ondubbelzinnig informeren van aankomende studenten over het opleidingsdoel en het beoordelingssysteem. Juist in opleidingen waarin de studenteninstroom divers is wat betreft hun beroepsperspectief (sommige studenten willen leraar worden, anderen komen vooral voor het vak en willen doorstromen naar universitair onderwijs), is een goede communicatie vereist over de consequenties van competentiegerichtheid. Er kan echter eveneens worden opgemerkt dat de EFA, door opleidingen zelf ruimte te geven voor experimenteren, ook een valkuil heeft weten te vermijden. Men heeft niet van boven af geprobeerd de eigen onzekerheid over assessment van competenties te overschreeuwen door op voorhand van alles vast te leggen, zoals de inrichting van het portfoliosysteem en het integraal beoordelingsmoment aan het einde van elke studiefase. Bepaald werd dat opleidingen met deze instrumenten zouden gaan werken, maar niet hoe zij dat moesten doen. Te veel vastleggen roept naar verwachting minstens zo veel tegenzin op als het onvermogen om, na een periode van experimenteren, gezamenlijke afspraken te maken waarin iedere opleiding zich kan vinden. Dit leidde tot het hybride beoordelingssysteem waarin studenten zowel kunnen worden toegelaten tot volgende studiefasen op basis van behaalde studiepunten (voor traditionele toetsen en practica) als op basis van een bewijs (met behulp van middelen zoals genoemd in figuur 2) dat hun competentieontwikkeling volgens een voorgeschreven tempo verloopt. In een dergelijk *hybride* systeem raken studenten al gauw in de verleiding om strategisch studiedrag te ontwikkelen en telkens die weg in te slaan die hen de beste kans op slagen lijkt te geven met de minste inspanning.

Twijfels over uitvoerbaarheid

Een tweede onderwerp waarmee men tijdens het traject naar competentiegericht beoordelen onvermijdelijk te maken krijgt, is de twijfel over de uitvoerbaarheid van het nieuwe beoordelingssysteem. Opleiders zijn in het denken over beoordeling geneigd om op 'safe' te spelen en traditionele toetsvormen te blijven gebruiken zolang nieuwe hun waarde nog onvoldoende bewezen hebben (zie ook Brown en Glasner, 1999). Zo ontstaat (tijdelijk) voor de studenten een opeenstapeling van toetsmomenten. Wij zijn van mening dat twijfel over de uitvoerbaarheid dikwijls te gemakkelijk wordt vertaald in de conclusie dat het doorvoeren van de innovatie praktisch onmogelijk is. Omgekeerd is het natuurlijk ook niet goed als twijfel wordt genegeerd of afgedaan met de aanbeveling om het nu eerst maar eens een tijdlang te proberen. De beste strategie is twijfel zorgvuldig in kaart te brengen. Voor sommige problemen is wellicht een betrekkelijk eenvoudige

oplossing te vinden. Er kan bijvoorbeeld worden afgesproken dat studenten die scoren op de grens tussen voldoende en onvoldoende voor een verlengde beoordelingsprocedure mogen kiezen. Een ander voorbeeld is de al eerder genoemde voortgangstoets. Het invoeren van een dergelijke toets is een goed antwoord op de twijfel die veel docenten hebben omtrent het portfolio en de kritische situaties als middelen om de expertkennis van studenten te beoordelen. In gevallen waarin twijfel niet op korte termijn is weg te nemen, kan het een goede strategie zijn om taken te splitsen en een duidelijk onderscheid te maken tussen de verschillende actoren die een rol spelen bij het assessment, zoals de beoordelingsautoriteit die verantwoordelijk is voor de inrichting van het assessment, de beoordelaars die het assessment uitvoeren en de mentoren die kandidaten begeleiden en adviseren bij het samenstellen van het portfolio. Ook het instellen van kwaliteitsbewakers in de vorm van 'examencommissies' die eventuele klachten van studenten afhandelen, is aan te raden (Hofstee, 1999; Elshout-Mohr et al., 2001).

Hoge eisen ten aanzien van elk afzonderlijk beoordelingsinstrument

Een verwant probleem aan het voorgaande is dat opleiders het standpunt innemen dat elk afzonderlijk beoordelingsinstrument aan de hoogste eisen van betrouwbaarheid en validiteit zou moeten voldoen. Soms wordt bijvoorbeeld gesteld dat kritische situaties als beoordelingsinstrument ongeschikt zouden zijn omdat het nu eenmaal onmogelijk is om kerncompetenties (of welke complexe vaardigheden dan ook) betrouwbaar te meten met behulp van een klein aantal kritische situaties. Deze stelling is uiteraard waar, maar impliceert niet dat kritische situaties geen adequaat *onderdeel* zouden kunnen zijn van een meeromvattend beoordelingssysteem. In figuur 2 zijn de kritische situaties waar assessoren dicht boven op zitten slechts een aanvulling op kritische praktijksituaties waarin de student is geobserveerd en beoordeeld door anderen. Wil men een uitspraak kunnen doen over de betrouwbaarheid of validiteit van nieuwe beoordelingsinstrumenten, dan zal men deze in hun context moeten onderzoeken. Dat wil zeggen: niet onafhankelijk van elkaar, maar in de samenhang waarin ze in de opleiding worden gebruikt. Wil men een nieuw beoordelingssysteem vergelijken met een traditioneel systeem, dan geldt voor het traditionele systeem hetzelfde. Een grote betrouwbaarheid en validiteit van bepaalde onderdelen van zo'n systeem impliceert nog niet dat het systeem als geheel aan dezelfde hoge standaarden voldoet. In feite geldt het bovenstaande ook voor pogingen om nieuwe instrumenten te verbeteren of te vervolmaken. Denk bijvoorbeeld aan het ontwikkelen van procedures om de interbeoordelaarsbe-

trouwbaarheid van 'peer-assessment' te vergroten. Het belang van dergelijke verbeteringen en de mate waarin deze een investering waard is, dient te worden gezien binnen het beoordelingssysteem als geheel en het succes ervan dient ook binnen dit geheel te worden geëvalueerd.

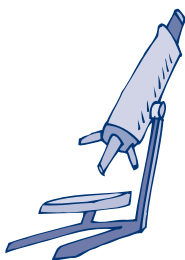
CONCLUSIE

Instellingen die overgaan naar competentiegericht leren, kunnen kiezen uit verschillende varianten. Naarmate de ontwikkeling van de beroepsidentiteit van de studenten een meer centrale plaats inneemt in het opleidingsconcept, zijn meer aanpassingen in het beoordelingssysteem noodzakelijk. Deze aanpassingen zijn op dit moment nog onvolledig beschreven en uitgewerkt, hetgeen de overstap naar competentiegericht opleiden tot een uitdaging maakt die alle kenmerken heeft van een innovatie. Een gedegen oriëntatie vooraf is noodzakelijk en strikte doordenking van de implicaties van het congruentieprincipe kan hierbij behulpzaam zijn. Een belangrijke keuze is of men het begrip 'competentiegericht' breed (gericht op beroepsvaardigheden en ontwikkeling van beroepsidentiteit) of smal (alleen gericht op beroepsvaardigheden) wil opvatten en of men daarbij denkt aan de lange of korte termijn. Gezien de complexiteit van de innovatie verdient het aanbeveling om van meet af aan samenwerking te zoeken met onderzoekers. Flankerend onderzoek kan ten dienste staan van ontwerpprocessen, het verwerven van systematische informatie omtrent reacties van docenten en studenten en versterkt de aansluiting met overeenkomstige ontwikkeltrajecten elders.



Literatuur

- Adviescommissie Onderwijs-Arbeidsmarkt, *De ontwikkeling van de kwalificatiestructuur voor het secundair beroepsonderwijs*. 's-Hertogenbosch: ACOA, 1997.
- Biggs, J.B., *Teaching for quality learning at university*. Open University Press, 1999.
- Black, P. en D. William, *Theory and practice in the development of formative assessment*. Keynote address EARLI-conference, 2001.
- Brown, S. en A. Glasner (eds.), *Assessment matters in higher education: Choosing and using diverse approaches*. Suffolk: St Edmundsbury Press, 1999.
- Dochy, F., M. Segers en D. Sluijsmans, 'The use of self-, peer-, and co-assessment in higher education: A review'. *Studies in Higher Education*, 24 (3), pp. 331-350, 1999.
- Driessen, E. en C. van der Vleuten, 'Matching student assessment to problem-based learning: lessons from experience in a law faculty'. *Studies in Continuing Education*, 22, pp. 253-249, 2000.
- Elshout-Mohr, M. en R. Oostdam, *Assessment van competenties in een dynamisch curriculum*. Amsterdam: SCO-Kohnstamm Instituut, 2001.
- Elshout-Mohr, M., R. Oostdam en M. Overmaat, 'Assessment van beroepscompetenties: Beslissingen bij het invoeren van integratieve beoordelingen in de opleiding'. *Tijdschrift voor Hoger Onderwijs*, 19 (3), pp. 205-221, 2001.
- Elshout-Mohr, M., R. Oostdam en M. Overmaat, 'Student assessment within the context of constructivistic educational settings'. *Studies in Educational Evaluation*, in druk.
- Elshout, M., R. Oostdam, A. Dietze en M. Snoek, 'Assessment van competenties'. *Tijdschrift voor lerarenopleiders*, 22 (2), pp. 48-54, 2001.
- Hofstee, W.K.B., *Principes van beoordelen*. Lisse: Swets & Zeitlinger, 1999.
- Korthagen, F., B. Koster, K. Melief en A. Tigchelaar, *Docenten leren reflecteren*. Soest: Uitgeverij Nelissen, 2002.
- Melief, K., A. Tigchelaar, F. Korthagen en B. Koster, *Leren van lesgeven*. Soest: Uitgeverij Nelissen, 2002.
- Mulder, M., *Competentieontwikkeling in organisaties*. 's-Gravenhage: Elsevier, 2002.
- O'Neil, Jr. H.F. (ed.), *Workforce Readiness*. Mahwah: Erlbaum, 1997.
- Overmaat, M., A. Dietze, I. Oudkerk Pool en A. Wilschut, 'Integratief beoordelen: een stap naar competentiegericht opleiden'. *VELON, Tijdschrift voor Lerarenopleiders*, 22 (2), pp. 13-20, 2001.
- Peetsma, T.T.D. en M.L.L. Volman (red.), *Bevorderende en belemmerende factoren voor een leven lang leren*. Amsterdam: SCO-Kohnstamm Instituut, 1997.
- Sandberg, J., A. Anjewierden, J. Groothuismink, R. de Hoog en M. Giebels, 'The CREDIT-project: Towards an international framework for assessment and accreditation'. *British Journal of Educational Technology*, 31 (4), pp. 281-298, 1999.
- Snoek, M., 'Aardverschuivingen in de lerarenopleidingen?' *VELON, Tijdschrift voor Lerarenopleiders*, 21 (3), pp. 5-12, 2000.
- Verloop, N., 'De leraar'. In: J. Lowyck en N. Verloop (red.), *Onderwijskunde: Een kennisbasis voor professionals* (pp. 109-150). Groningen: Wolters-Noordhoff, 1995.
- Wielenga, D., *Probing and proving competence*. Paper gepresenteerd op de conferentie 'The new Educational Benefits of ICT. ICT in Higher Education'. Te vinden onder <<http://www.efa.nl/publicaties/>>, 2002.



360°-FEEDBACK ALS HULPMIDDEL VOOR HET EVALUEREN VAN OPLEIDING EN ONTWIKKELING BINNEN BEDRIJFS- ORGANISATIES

FROUKJE JELLEMA

9

Dit artikel gaat in op evaluatie in de context van bedrijfsopleidingen. Organisaties hebben in toenemende mate behoefte aan inzicht in de effecten van opleiding en ontwikkeling van werknemers, vooral ten aanzien van veranderd werkgedrag. Als gevolg hiervan valt er een behoefte waar te nemen aan evaluatiemethoden die zich richten op veranderd werkgedrag. Er is echter een tekort aan praktische en bruikbare methoden. Dit maakt het bestuderen van 360°-feedback als mogelijke methode interessant. Bij 360°-feedback ontvangt een medewerker feedback over het werkgedrag van een aantal mensen uit de werkomgeving, zoals de leidinggevende, collega's en medewerkers.

INLEIDING

De belangrijkste vraag die in dit artikel wordt gesteld is: Is 360°-feedback een geschikte methode voor het evalueren van veranderd werkgedrag als gevolg van opleiding en ontwikkeling? Deze vraag wordt beantwoord aan de hand van vier deelvragen:

1. Wat is een verklaring voor de belangstelling in organisaties voor evaluatie van werkgedrag?
2. Hoe wordt 360°-feedback in Nederland voornamelijk gebruikt?

3. Wat zijn voorwaarden voor het deugdelijk gebruik van 360°-feedback en in hoeverre wordt hieraan voldaan in Nederlandse organisaties? Bij deze vraag gaat het nog om het gebruik van 360°-feedback in het algemeen, los van de toepassing 'evaluatie'.
4. Hoe kan 360°-feedback worden gebruikt voor opleidingsevaluatie?

Eerst wordt kort ingegaan op mogelijke effecten van opleidingen en mogelijke meetmethoden. Daarna worden de vier bovengenoemde vragen beantwoord. Overigens wordt in dit artikel voornamelijk gesproken over *bedrijfsopleidingen*, alhoewel dit onderwerp vanzelfsprekend ook relevant is voor andere soorten ontwikkelingsinterventies.

HET EVALUEREN VAN VERANDERD WERKGEDRAG ALS GEVOLG VAN OPLEIDING

Er wordt veel geld en tijd geïnvesteerd in het opleiden van werknemers. Los van de soort opleiding (bijvoorbeeld een cursus buiten het bedrijf of leren op de werkplek) hebben organisaties behoefte aan inzicht in de effectiviteit van opleidingsinspanningen. Een

belangrijk aspect waarop opleidingen kunnen worden geëvalueerd is veranderd werkgedrag. Het gaat dan om de vraag of het geleerde daadwerkelijk wordt toegepast op de werkvloer (de 'transfer' van een opleiding). Evaluatie van werkgedrag vindt slechts in beperkte mate plaats. Volgens een Amerikaans onderzoek wordt werkgedrag voor maar 12 procent van de opleidingen geëvalueerd (Bassi en Van Buren, 1999). Het merendeel van de opleidingen (88 procent) wordt dus niet geëvalueerd wat betreft het werkgedrag. Redenen hiervoor zijn onder andere dat systematische evaluatie duur is, veel tijd kost en soms onnodig of bedreigend wordt gevonden. Bovendien beschikken organisaties vaak niet over praktische en bruikbare methoden.

Organisaties tonen echter steeds meer interesse in de evaluatie van opleidingen ten aanzien van werkgedrag. Wellicht dat dit het gevolg is van de toenemende onderkenning van een 'transferprobleem': slechts een klein percentage van de investeringen in opleidingen leidt daadwerkelijk tot transfer van het geleerde naar de werksituatie (McCarthy en Garavan, 1999). Een andere verklaring voor de toenemende belangstelling kan de opkomst van het 'competentiedenken' zijn. Onder competenties wordt verstaan: 'het cluster van kennis, vaardigheden en attitudes die noodzakelijk zijn voor effectief functioneren in een bepaald beroep, organisatie, functie of rol' (Mulder, 2001). Competenties zijn niet hetzelfde als werkgedrag maar een voorwaarde voor het gedrag. Omdat competentie gezien wordt als een vermogen tot gedrag, kan competentie niet direct worden gemeten. Vandaar dat competenties meestal worden gemeten via beoordeling van het werkgedrag. Van oudsher is de leidinggevende de voornaamste beoordelaar van het werkgedrag van medewerkers. Aangezien er steeds meer in teams wordt gewerkt en organisaties platter worden, heeft de leidinggevende meer en meer een beperkt zicht op het werkgedrag van medewerkers (Bracken, 1994; Cascio, 1995; Murphy en Cleveland, 1995). Dat verklaart de opkomst in het gebruik van 360°-feedback, waarbij ook anderen dan de leidinggevende het werkgedrag beoordelen (Van Velsor et al., 1997).

Een aantal methoden om werkgedrag te evalueren zijn (Swanson en Holton, 1997):

- *Observatie.* Iemand in de omgeving van de deelnemer observeert het werkgedrag. Hierbij kan het gaan om observatie op de werkplek (het feitelijke gedrag) of om observatie buiten de werkplek (bijvoorbeeld beoordeling door experts in een gesimuleerde situatie).
- *Zelfbeoordeling.* Deelnemers aan de bedrijfsopleiding worden, bijvoorbeeld door middel van een vragen-

lijst, zelf verzocht aan te geven in hoeverre hun werkgedrag veranderd is als gevolg van de opleiding. Deze methode wordt erg veel gebruikt (Baldwin en Ford, 1988; Hazucha, Hezlett en Scheider, 1993). Soms wordt ook de leidinggevende van de deelnemers gevraagd een dergelijke vragenlijst in te vullen (McLean, Sytsma en Kerwin-Ryberg, 1995).

- *Objectieve methoden.* Hierbij wordt geprobeerd tot een zo accuraat mogelijke beoordeling van het werkgedrag te komen. Het kan dan bijvoorbeeld gaan om beoordeling met behulp van testen. Voor een groot aantal opleidingen is het mogelijk objectief te testen of een deelnemer de nieuw geleerde vaardigheden in de praktijk toepast, bijvoorbeeld voor allerlei technische opleidingen.

360°-FEEDBACK: HUIDIGE TOEPASSINGEN

Bij 360°-feedback krijgt een werknemer (in het vervolg de feedbackontvanger genoemd) feedback op het eigen functioneren van verschillende personen uit de werkomgeving, zoals de leidinggevende, collega's, medewerkers, klanten en leveranciers (in het vervolg feedbackgevers genoemd). Daarnaast geeft de feedbackontvanger zijn of haar eigen mening over het functioneren.

Op het moment wordt 360°-feedback vooral gebruikt ter ondersteuning van personeelsontwikkeling. Van de Nederlandse organisaties die 360°-feedback gebruiken, doet 93 procent dit in het kader van personeelsontwikkeling (Jellema, 2000a). De feedback is in dat geval aanleiding voor de feedbackontvanger om bepaalde competenties te ontwikkelen. De achterliggende gedachte is dat werknemers eerst inzicht in hun sterke punten en ontwikkelpunten moeten hebben voordat ze zich kunnen ontwikkelen (Borman, 1997; Garavan, Morley en Flynn, 1997). Een andere mogelijke toepassing is het gebruik van 360°-feedback als basis voor het nemen van beslissingen over salaris en promotie. In het algemeen overheerst echter in de literatuur de mening dat met deze toepassing van 360°-feedback zeer voorzichtig moet worden omgegaan, aangezien dit veel onzekerheid en weerstand kan oproepen. Een aantal onderzoeken laat bijvoorbeeld zien dat feedbackgevers positievere feedback geven wanneer de feedback voor een dergelijke toepassing wordt gebruikt (London en Beatty, 1993). Andere mogelijke toepassingen van 360°-feedback zijn het ondersteunen van cultuurverandering, stimuleren van communicatie, versterken van de kernwaarden van de organisatie en blootleggen en oplossen van conflicten.

VOORWAARDEN VOOR DEUGDELIJK GEBRUIK VAN 360°-FEEDBACK

Als 360°-feedback wordt gebruikt, is het van belang voldoende aandacht te geven aan het invoeringsproces en de keuze van de instrumenten. Beide aspecten worden hierna besproken.

Voorwaarden voor het 360°-feedbackproces: literatuur en Nederlandse praktijk

Het 360°-feedbackproces moet zorgvuldig plaatsvinden. Op grond van literatuuronderzoek kunnen de volgende aandachtspunten worden bepaald:

1. Ten eerste is het van belang dat alle betrokkenen voldoende worden geïnformeerd. De informatie betreft dan de toepassing waarvoor de 360°-feedback gebruikt gaat worden, namelijk opleidings-evaluatie, alsmede de manier waarop 360°-feedback wordt verzameld, verwerkt en teruggekoppeld.
2. Ook de selectie van feedbackgevers is belangrijk. De feedbackgevers moeten in de positie zijn om het werkgedrag van de feedbackontvanger zowel voor als na de opleiding te observeren en te beoordelen. De kans hierop is het grootst als de feedbackontvanger zelf de feedbackgevers selecteert, omdat de feedbackontvanger het beste weet met wie hij of zij regelmatig samenwerkt. Een nadeel hiervan kan zijn dat de feedbackontvanger enkel mensen uitkiest waarvan hij of zij positieve feedback verwacht, maar dit blijkt in de praktijk niet een groot probleem te zijn (Edwards en Ewen, 1996; Fletcher en Baldry, 1999). Eventueel kan de feedbackontvanger de feedbackgevers samen met de leidinggevende selecteren.
3. Om betrouwbaarheid te stimuleren wordt het in de literatuur belangrijk geacht dat feedbackgevers anoniem hun feedback kunnen geven. De aanname is dat in dat geval de feedback eerlijker en daarmee betrouwbaarder is. Soms wordt in de literatuur ook wel gesteld dat de betrouwbaarheid toeneemt als er juist geen sprake is van anonimiteit. De feedbackgevers weten dan namelijk dat ze kunnen worden aangesproken op hun feedback. Als de organisatie echter weinig ervaring heeft met 360°-feedback, is het beter om in het begin de anonimiteit te garanderen. Dit kan bijvoorbeeld door voldoende feedbackgevers erbij te betrekken zodat de oordelen kunnen worden gemiddeld.
4. Uit de literatuur blijkt dat het ook belangrijk is om feedbackgevers van tevoren te trainen in het gebruik van het 360°-feedbackinstrument (meestal een vragenlijst) en het beoordelen van werkgedrag.
5. Ten slotte, gezien de positieve invloed van 360°-feedback op de persoonlijke ontwikkeling, moeten de resultaten worden teruggekoppeld naar de feedbackontvanger, ook al is dit in het geval van gebruik

van 360°-feedback voor opleidingsevaluatie niet de belangrijkste toepassing. Als resultaten niet worden teruggekoppeld en de feedbackontvangers niet de kans krijgen van de feedback te leren, zal de acceptatie en deelname aan de evaluatie gering zijn.

In het algemeen blijkt dat in Nederlandse organisaties redelijk aan de voorwaarden voor het feedbackproces wordt voldaan (Jellema, 2000a). Communicatie over het doel van de feedback vindt in vrijwel alle gevallen plaats en feedbackontvangers hebben meestal invloed op de keuze van de feedbackgevers. Aangezien feedbackontvangers vaak (mede) bepalen wie feedback gaan geven, mag worden aangenomen dat de geselecteerde feedbackgevers in de positie zijn het gedrag van de feedbackontvanger te observeren en te beoordelen. De follow-up is in de meeste gevallen voldoende: feedbackontvangers krijgen niet alleen hulp bij interpretatie en verwerking van de feedback, maar ook de gelegenheid om aan follow-up-activiteiten deel te nemen. Minder positief aan het feedbackproces in Nederlandse organisaties is dat feedbackgevers in het algemeen onvoldoende training krijgen in het beoordelen en geven van feedback, terwijl dit in verband met de betrouwbaarheid van de beoordelingen wel belangrijk is. Bovendien is de vertrouwelijkheid van de feedback niet altijd voldoende gewaarborgd: hoewel de feedback in de meeste gevallen anoniem wordt gegeven, zijn er niet altijd voldoende feedbackgevers bij betrokken om vertrouwelijkheid te kunnen garanderen.

Voorwaarden voor 360°-feedbackinstrumenten: literatuur en Nederlandse praktijk

Ook het 360°-feedbackinstrument dat wordt gebruikt is van belang. In de literatuur worden enkele belangrijke instrumentkenmerken onderscheiden:

1. Een eerste belangrijk kenmerk is de inhoud van het gebruikte instrument. Meestal bestaan 360°-feedbackinstrumenten uit een aantal overkoepelende categorieën gedragskenmerken of competenties, die op hun beurt bestaan uit een aantal items. Om de betrouwbaarheid van de beoordeling te garanderen worden veelal gedragsgeoriënteerde items gebruikt (Dalessio, 1998; Yukl en Lepsinger, 1995).
2. Van belang is verder dat het instrument alleen items bevat die ook daadwerkelijk ontwikkelbaar zijn en kunnen worden beïnvloed door de feedbackontvanger. Omdat de competenties en items die in het 360°-feedbackinstrument worden opgenomen, relevant moeten zijn voor de opleiding die wordt geëvalueerd, is het zinvol de opleider bij de selectie te betrekken.
3. Een ander punt is de antwoordschaal die wordt gebruikt. Aangezien het doel is om ontwikkeling in werkgedrag te meten, moet de antwoordschaal

hiervoor wel voldoende ruimte bieden. Vaak wordt gesteld dat een 5-puntsschaal afdoende is, maar ook een 7-punts- of 8-puntsschaal kan een optie zijn.

4. Ook de psychometrische kenmerken van het gebruikte instrument, de betrouwbaarheid en de validiteit, zijn belangrijk. De betrouwbaarheid en validiteit van 360°-feedbackinstrumenten worden hierna wat uitgebreider besproken.

Betrouwbaarheid

De betrouwbaarheid van 360°-feedbackinstrumenten kent verschillende aspecten (Dalessio, 1998; Van Velsor, Leslie en Fleenor, 1997), te weten de interne consistentie, test-hertest betrouwbaarheid en interbeoordelaarsbetrouwbaarheid.

Bij interne consistentie gaat het om de bijdrage van elk item aan de bijbehorende competentie, meestal gemeten met Cronbachs alfa. Voordat een instrument wordt ingezet, moet er duidelijkheid zijn over de interne consistentie van de gebruikte schalen.

Test-hertest betrouwbaarheid betreft de overeenstemming in de beoordeling van dezelfde feedbackontvanger door dezelfde feedbackgever op twee verschillende meetmomenten. Wanneer het instrument op twee verschillende momenten wordt gebruikt, zonder dat er in de tussentijd iets is veranderd, zouden de oordelen op de twee meetmomenten in principe moeten overeenstemmen.

Bij interbeoordelaarsbetrouwbaarheid gaat het om de overeenstemming tussen de oordelen van verschillende feedbackgevers. Over interbeoordelaarsbetrouwbaarheid van 360°-feedbackinstrumenten bestaat veel discussie. Deze instrumenten worden juist gebruikt om verschillende perspectieven op het gedrag te achterhalen. Een lage betrouwbaarheid kan dus worden opgevat als een argument voor het gebruik van deze instrumenten. Als de interbeoordelaarsbetrouwbaarheid hoog is, zou dit betekenen dat kan worden volstaan met één feedbackgever (bijvoorbeeld de leidinggevende). Echter, als er sprake is van zorgvuldige itemformulering, zou er een zekere overeenstemming moeten zijn. Bij beoordelaarsbetrouwbaarheid kan een onderscheid worden gemaakt tussen betrouwbaarheid binnen groepen (bijvoorbeeld collega's onderling of medewerkers onderling) en betrouwbaarheid tussen groepen (bijvoorbeeld tussen collega's en medewerkers). Vaak wordt gesteld dat de betrouwbaarheid binnen groepen hoger moet zijn dan die tussen groepen. Als 360°-feedback wordt gebruikt voor opleidingsevaluatie, is het aggregatieniveau waarop de resultaten worden gepresenteerd van belang. Er zijn diverse mogelijkheden: een score per individuele feedbackgever, een gemiddelde score per feedbackgeversgroep of een gemiddelde van alle feedbackgevers tezamen.

Als de interbeoordelaarsbetrouwbaarheid hoog is, zouden de resultaten als één gemiddelde score kunnen worden gepresenteerd; bij een lage beoordelaarsbetrouwbaarheid kunnen de resultaten daarentegen beter per feedbackgeversgroep worden gepresenteerd. Als de betrouwbaarheid binnen een groep ook laag is, zouden de resultaten in principe ook nog per individuele feedbackgever kunnen worden gepresenteerd. Vanwege de anonimiteit wordt dit laatste echter vrijwel niet gedaan.

Validiteit

Over de vraag of 360°-feedbackinstrumenten valide moeten zijn (met andere woorden: meet het instrument wat het zou moeten meten?), verschillen de meningen. Alhoewel een beperkt aantal auteurs de validiteit van instrumenten heeft onderzocht (Church, 1999; McEnery en Blanchard, 1999), wordt door de bureaus die de instrumenten leveren vaak geopperd dat 360°-feedback nou eenmaal een subjectief beeld geeft van het gedrag van de feedbackontvanger. Zelfs als dit beeld niet geheel overeenstemt met het werkelijke (objectieve) gedrag, is het toch vaak de basis voor personeelsbeslissingen. Ook vanuit de wetenschap wordt wel eens geopperd dat validiteit van competentiemetingen niet van belang is (zie bijvoorbeeld het interview met professor Mulder in *PW Vakblad voor Personeelsmanagement*, nummer 7, 2002). Toch is het voor zowel wetenschap als praktijk van belang om onderzoek te doen naar de validiteit van dergelijke instrumenten. Immers, wanneer de resultaten van de 360°-feedback de basis zijn voor ingrijpende beslissingen, bijvoorbeeld een bepaalde feedbackontvanger promotie geven of een bepaalde opleiding afstoten, moeten er wel aanwijzingen zijn dat de informatiebasis valide is. Om meer te weten te komen over de validiteit van 360°-feedback kan het zinvol zijn de evaluatie met 360°-feedback aan te vullen met een andere, objectievere evaluatiemethode. Bovendien kunnen naast de 360°-feedback ook achtergrondgegevens in de evaluatie worden meegenomen, zoals motivatie van de deelnemer en ondersteuning op de werkplek, om te helpen bij de interpretatie van de evaluatieresultaten.

Beperkt onderzoek

Zoals hiervoor al geschetst is, blijkt in het algemeen dat Nederlandse 360°-feedbackinstrumenten vooral ontwikkeld zijn ten behoeve van ondersteuning van de persoonlijke ontwikkeling (Jellema, 2000b). De meeste instrumenten bevatten een aantal overkoepelende categorieën competenties, waarbij er vrijwel altijd sprake is van een uitgebreid databestand waaruit de klant kan kiezen. In het algemeen lijkt er een trend te zijn in de richting van meer maatwerk. De antwoordschalen van de bestudeerde instrumenten lopen uiteen van 4-puntsschalen tot 9-puntsschalen en richten zich voor-

al op de mate waarin men het met een stelling eens is of op de frequentie waarin bepaald gedrag wordt vertoond. Onderzoek naar de psychometrische kenmerken van Nederlandse instrumenten is beperkt. Slechts een klein aantal instrumenten is onderzocht op betrouwbaarheid. Meestal betreft het dan onderzoek naar de interne consistentie, die vaak voldoende is, en de interbeoordelaarsbetrouwbaarheid, die vaak matig is. Onderzoek naar de validiteit van de instrumenten wordt vrijwel niet gedaan.

HOE KAN 360°-FEEDBACK WORDEN GEBRUIKT VOOR OPLEIDINGSEVALUATIE?

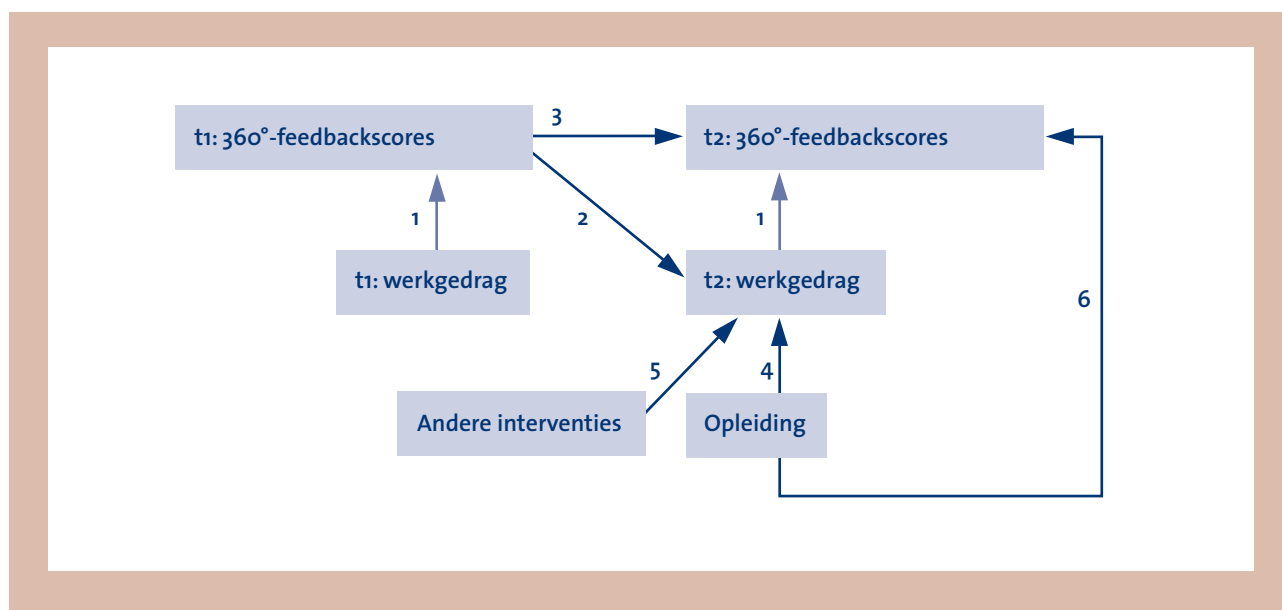
Hierna wordt eerst een model voor opleidingsevaluatie met 360°-feedback besproken, gevolgd door een mogelijke methode om een dergelijke evaluatie uit te voeren.

Een model voor opleidingsevaluatie met 360°-feedback

Alhoewel in diverse evaluatiehandboeken wordt geopperd dat 360°-feedback een geschikte methode is voor evaluatie van bedrijfsopleidingen (zie bijvoorbeeld Brown en Seidner, 1998; Kirkpatrick, 1994; Robinson en Robinson, 1989), is hier slechts in beperkte mate onderzoek naar gedaan (uitzonderingen zijn bijvoorbeeld McLean, Sytsma en Kerwin-Ryberg, 1995; Rosti en Shipper, 1998; Van Sandick en Schaap-Neuteboom, 1993). Opleidingsevaluatie met 360°-feedback lijkt relatief simpel (Fletcher en Baldry, 1999): feedback wordt zowel voor als na de interventie verzameld en het verschil tussen de t1-scores (voor de opleiding) en

t2-scores (na de opleiding) representeert het opleidingseffect. In de praktijk is dit echter een stuk ingewikkelder omdat er een aantal beïnvloedende factoren zijn. In het kader van een promotieonderzoek naar evaluatie met 360°-feedback is een model ontwikkeld (zie figuur 1) waarin de beïnvloedende factoren zijn weergegeven (Jellema, in druk). In het model worden overigens twee verschillende soorten pijlen gebruikt: pijl 1 vertegenwoordigt een operationalisering, terwijl de pijlen 2 tot en met 6 een causaal verband weergeven. Het model wordt hierna toegelicht.

Als eerste kan worden gesteld dat het meten van werkgedrag erg moeilijk is omdat werkgedrag door veel factoren wordt beïnvloed. Zeker als het werkgedrag wordt gemeten door middel van oordelen van anderen, kunnen veel factoren de beoordeling beïnvloeden, zoals het doel van de meting, het meetproces, het meetinstrument, kenmerken van de feedbackgever en kenmerken van de feedbackontvanger (Landy en Farr, 1983). Bij 360°-feedback wordt het werkgedrag niet direct gemeten maar door middel van de oordelen van mensen in de werkomgeving van de feedbackontvanger (pijl 1 in figuur 1). Dit levert meteen de eerste moeilijkheid op. Ook al is zowel het werkgedrag als de perceptie van mensen in de werkomgeving op dit werkgedrag van belang voor de loopbaan van medewerkers, het is belangrijk te onderkennen dat het twee verschillende zaken zijn. Het gebruik van 360°-feedback impliceert bovendien de betrokkenheid van meerdere beoordelaars. Het feit dat er verschillende feedbackgevers betrokken zijn bij het proces kan enerzijds als een aanvulling worden gezien, maar roept anderzijds de vraag op wat met een eventuele lage interbeoordelaarsbetrouwbaarheid moet worden gedaan.



Figuur 1. Beïnvloedende factoren op opleidingsevaluatie met 360°-feedback (Jellema, in druk)

Een tweede beïnvloedende factor is dat 360°-feedback oorspronkelijk bedoeld is als ontwikkelingsinstrument en niet als evaluatie-instrument. Dit betekent dat de feedback die voorafgaand aan de opleiding wordt gegeven het werkgedrag na de opleiding kan beïnvloeden (pijl 2 in de figuur). Sterker nog: dit is vaak de belangrijkste doelstelling van 360°-feedback. Uit een meta-analyse naar het effect van feedback (Kluger en DeNisi, 1996) blijkt dat feedback een groot effect heeft op de prestaties van werknemers. Het zal dus duidelijk zijn dat wanneer 360°-feedback als evaluatiemethode wordt gebruikt, twee interventies door elkaar lopen: enerzijds wordt een effect van de opleiding verwacht, maar anderzijds heeft de 360°-feedback zelf ook effecten.

Een derde factor is dat de feedback op t1 ook de feedback op t2 kan beïnvloeden, los van de invloed op het werkgedrag (pijl 3 in de figuur). Denk hierbij bijvoorbeeld aan het feit dat veel feedbackontvangers ‘overschatters’ zijn (Yammarino en Atwater, 1997). Dit betekent dat op tijdstip t1 de zelfscores hoger zijn dan de scores van de andere feedbackgevers. Vaak passen feedbackontvangers daarom hun zelfscores aan bij een volgende meting. Feedback op t2 kan ook worden beïnvloed door herhaaldelijk gebruik van het instrument. Denk bijvoorbeeld aan het ‘plafondeffect’: als scores bij de eerste meting erg hoog zijn, is er weinig ruimte voor hogere scores bij de tweede meting.

Behalve de opleiding (pijl 4) kunnen ook andere interventies het werkgedrag beïnvloeden (pijl 5). Vooral wanneer er sprake is van een controlegroep naast een experimentele groep (een beschrijving van het evaluatieontwerp volgt later in de tekst), kan dit een moeilijk punt zijn. Voor de evaluatie is het immers onmogelijk om betrokkenen in de controlegroep te weerhouden van ontwikkelingsactiviteiten. Daarnaast kan het werkgedrag ook worden beïnvloed door allerlei veranderingen in de context, zoals een nieuwe leidinggevende, ontwikkelingen op de markt of nieuw beleid. Daarom wordt het door veel auteurs van belang geacht opleidingseffecten zorgvuldig te isoleren van eventuele andere effecten, alhoewel dat tegelijkertijd als extreem moeilijk wordt gezien.

Een laatste factor die de evaluatie kan beïnvloeden is dat de interventie niet alleen het t2-werkgedrag beïnvloedt, maar ook de t2-feedback los van het werkgedrag (pijl 6). Relevant in deze context is het onderscheid tussen alfa-, bèta- en gammaverandering (Golembiewski, Billingsley en Yeager, 1976). Alfaverandering betekent dat het werkgedrag verandert binnen een relatief stabiel systeem. Bèta-verandering daarentegen betekent dat de context waarin de beoordeling plaatsvindt veranderd kan zijn. Bijvoorbeeld: doordat een medewerker naar een opleiding

gaat, kunnen mensen in de werkomgeving andere verwachtingen van het werkgedrag krijgen. Bij gammaverandering is sprake van een herdefinitie van de relevante concepten die worden beoordeeld, wat vaak juist een belangrijk doel van opleiden is. Een voorbeeld is een werknemer die er door deelname aan een communicatieopleiding achter komt dat er veel meer bij communiceren komt kijken dan hij of zij dacht en zichzelf bij de nameting negatiever beoordeelt dan bij de voormeting. Omdat beide beoordelingen vanuit een ander referentiekader plaatsvinden, zijn ze moeilijk met elkaar te vergelijken.

Een methode voor opleidingsevaluatie met 360°-feedback

De wijze waarop 360°-feedback wordt gebruikt voor evaluatie moet zo veel mogelijk tegemoetkomen aan de beïnvloedende factoren die hiervoor genoemd zijn. Naast het 360°-feedbackproces en het gebruikte instrument is ook het evaluatieontwerp van belang. In de literatuur worden enkele mogelijke ontwerpen genoemd (zie Robinson en Robinson, 1989). Een eerste mogelijkheid is bijvoorbeeld het ‘one-shot’ evaluatieontwerp, waarbij eenmalig na de opleiding wordt gemeten of de deelnemers een bepaald niveau hebben bereikt. Dit is het meest voorkomende ontwerp. Een ander evaluatieontwerp is een ontwerp waarbij een voor- en een nameting worden afgenomen bij de deelnemers van de opleiding. Op deze manier is het mogelijk verandering te meten. Het belangrijkste nadeel van dit ontwerp is dat het niet mogelijk is opleidingseffecten te isoleren aangezien er geen vergelijkbare groep is die niet naar de opleiding is geweest. Een andere optie is daarom het trained/untrained design. Hierbij wordt gebruikgemaakt van een voor- en nameting met 360°-feedback bij zowel een experimentele groep als een controlegroep. De effecten van de opleiding zijn dan het verschil in vooruitgang tussen de experimentele groep en de controlegroep. Bij voorkeur worden dezelfde feedbackgevers bij de voor- en nameting betrokken. Omdat soms wordt gesteld dat de feedback erg afhankelijk is van de individuele feedbackgever, is dit in eerste instantie de beste manier om verandering te meten. Zoals al gezegd, moet aandacht worden besteed aan de mogelijkheid van bèta- of gammaverandering. Een oplossing hiervoor is de nameting uit te breiden met een retrospectieve vragenlijst, waarin de deelnemer gevraagd wordt aan te geven hoe hij of zij achteraf denkt over het functioneren voorafgaand aan de opleiding. Een ander alternatief is, na de opleiding, direct te vragen naar geobserveerde verandering in het werkgedrag van de feedbackontvanger.

In tabel 1 wordt de manier waarop opleidingsevaluatie met 360°-feedback kan plaatsvinden samengevat.

ASPECT	AANBEVELINGEN
360°-feedback-proces	Informeer betrokkenen over het doel en de procedure van de 360°-feedback. Laat de feedbackontvanger zelf de feedbackgevers selecteren. Vraag dezelfde feedbackgevers bij de voor- en nameting. Vraag voldoende feedbackgevers om anonimiteit en betrouwbaarheid te kunnen garanderen. Zorg ervoor dat feedbackgevers de feedback anoniem kunnen geven. Train de feedbackgevers van tevoren in het gebruik van de vragenlijst en het beoordelen.
360°-feedback-instrument	Koppel de 360°-feedbackresultaten terug naar de deelnemers (bespreken, follow-up). Betrek de trainer bij de selectie van competenties in het instrument. Gebruik een antwoordschaal met voldoende punten om verandering te meten. Gebruik een instrument met voldoende betrouwbaarheid. Is interbeoordelaarsbetrouwbaarheid voldoende, presenteer de resultaten dan als een gemiddelde score; is de interbeoordelaarsbetrouwbaarheid onvoldoende, presenteer de resultaten dan per feedbackgeversgroep. Vergelijk de 360°-feedback met een andere, wellicht objectievere methode om validiteit te bestuderen.
Evaluatieontwerp	Gebruik een aanvullende vragenlijst over achtergrondgegevens en contextfactoren. Gebruik een experimentele groep en een controlegroep. Gebruik een voor- en een nameting. Gebruik een retrospectieve voormeting om het effect van bèta- en gammaverandering dan wel het effect van t1-feedback te ondervangen.

Tabel 1. Opleidingsevaluatie met 360°-feedback

De mening van experts over opleidingsevaluatie met 360°-feedback

De vraag of 360°-feedback een geschikte methode is voor opleidingsevaluatie, is voorgelegd aan een aantal experts op het gebied van 360°-feedback en evaluatie en aan een aantal opleidingsmanagers. Alhoewel de respondenten in het algemeen aangeven dat 360°-feedback een geschikte methode is voor het evalueren van bedrijfsopleidingen, plaatsen zij er ook een aantal kanttekeningen bij. Deze kanttekeningen vallen grofweg uiteen in twee categorieën. Enerzijds zijn er respondenten die veel belang hechten aan het gebruikte instrument en stellen dat deze doorgaans niet in staat zijn het werkgedrag voldoende betrouwbaar en valide te meten. Daarnaast zijn er ook respondenten die het gebruikte instrument juist niet zo belangrijk vinden, maar die denken dat het 360°-feedbackproces en het evaluatieontwerp in de praktijk moeilijk te organiseren zijn.

De experts hebben ook hun mening gegeven over de voorgestelde methode. Niet alle aanbevelingen werden door hen even belangrijk geacht. Vooral anonimiteit van de feedbackgever werd niet zo belangrijk gevonden. Een aantal experts was van mening dat feedback die niet anoniem wordt gegeven juist betrouwbaarder is en bovendien waardevoller voor de feedbackontvanger. De selectie van de competenties in het instrument kan volgens de meeste experts beter overgelaten

worden aan de organisatie zelf in plaats van aan de trainer. Het meest opvallende resultaat was dat een controlegroep niet altijd belangrijk wordt gevonden voor evaluatie. Vooral vanuit de praktijk wordt gesteld dat een controlegroep niet praktisch en realiseerbaar is.

Zoals gezegd is er nog maar weinig ervaring met het gebruik van 360°-feedback voor evaluatie van veranderd werkgedrag. In het kader van het al eerder vermelde promotieonderzoek (Jellema, in druk) zijn evaluaties uitgevoerd binnen drie organisaties, te weten het Onze Lieve Vrouwe Gasthuis in Amsterdam, HBG en de gemeente Den Haag. Op grond van de resultaten kunnen enkele aanbevelingen worden gedaan:

- Zorg dat het instrument is afgestemd op de opleiding die wordt geëvalueerd en op de werksituatie van de feedbackontvangers. Als dit niet het geval is, geven feedbackgevers vaak aan dat ze het moeilijk vinden om feedback te geven. Een instrument dat niet relevant is kan zelfs weerstand oproepen.
- Aangezien de feedback erg blijkt af te hangen van de individuele feedbackgever, is het nog belangrijker dan al gedacht werd om dezelfde personen in de voor- en nameting te betrekken. Als de inschatting is dat medewerkers geen moeite hebben met het geven van feedback, kan daarnaast worden overwogen om de aanbeveling van anonimiteit te laten vallen. Dit heeft als voordeel dat per feedbackgever de feedback van de voormeting kan worden vergeleken met die van de nameting.
- 360°-feedback kan nooit enkel voor evaluatie worden gebruikt. De voordelen voor de ontwikkeling van

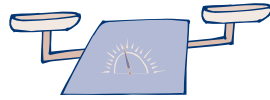
medewerkers zijn simpelweg te groot. De mogelijkheden voor feedbackontvangers om van de feedback te leren en zich verder te ontwikkelen, moeten worden benut.

Ten slotte: evaluatie van bedrijfsopleidingen met 360°-feedback is erg complex. Een alternatief is om de evaluatie niet aan een opleiding te koppelen, maar om medewerkers zelf de verantwoording te geven over het monitoren van hun ontwikkelingspanningen van de afgelopen periode. Als 360°-feedback bijvoorbeeld jaarlijks of tweejaarlijks wordt verzameld, krijgen medewerkers op die manier een goed inzicht in hun ontwikkeling.



Literatuur

- Baldwin, T.T. en J.K. Ford, 'Transfer of training: a review and directions for future research'. *Personnel Psychology*, 41 (1), pp. 63-105, 1988.
- Bassi, L.J. en M.E. van Buren, *The 1999 ASTD State of the Industry Report*. A supplement to *Training & Development Magazine*, 1999.
- Borman, W.C., '360-degree ratings: an analysis of assumptions and a research agenda for evaluating their validity'. *Human Resource Management Review*, 7 (3), pp. 299-315, 1997.
- Bracken, D.W., 'Straight talk about multirater feedback'. *Training & Development*, 48 (9), pp. 44-51, 1994.
- Brown, S.M. en C.J. Seidner, *Evaluating corporate training: models and issues*. Boston: Kluwer Academic Publishers, 1998.
- Cascio, W.F., *Managing human resources*. New York: McGraw-Hill, 1995.
- Church, A.H., *Do higher performing managers actually receive better ratings? A validation of multirater assessment methodology*. Academy of Human Resource Development, conference proceedings, 1999.
- Dalessio, A., 'Using multisource feedback for employee development and personnel decisions'. In: J.W. Smither (ed.), *Performance appraisal: state of the art in practice*. San Francisco: Jossey-Bass, 1998.
- Edwards, M.R. en A.J. Ewen, *360° feedback: the powerful new model for employee assessment & performance improvement*. New York: AMACOM, 1996.
- Fletcher, C. en C. Baldry, 'Multi-source feedback systems: a research perspective'. *International Review of Industrial and Organizational Psychology*, 14, pp. 149-193, 1999.
- Garavan, T., M. Morley en M. Flynn, '360 degree feedback: its role in employee development'. *Journal of Management Development*, 16 (2/3), pp. 134-147, 1997.
- Golembiewski, R.T., K. Billingsley en S. Yeager, 'Measuring change and persistence in human affairs: types of change generated by OD designs'. *The Journal of Applied Behavioral Science*, pp. 133-157, 1976.
- Hazucha, J., S. Hezlett en R. Schneider, 'The impact of 360-degree feedback on management skills development'. *Human Resource Management*, 32 (2/3), pp. 325-351, 1993.
- Jellema, F., 'Toepassing van 360-graden feedback in Nederlandse organisaties'. *Opleiding & Ontwikkeling*, (7/8), pp. 27-31, 2000a.
- Jellema, F., 'De kwaliteit van 360-graden feedback instrumenten: 7 Nederlandse instrumenten nader bekeken'. *M&O*, mei/juni, pp. 23-35, 2000b.
- Jellema, F. *Measuring training effects: the potential of 360-degree feedback*. Proefschrift, Universiteit Twente, in druk.
- Kirkpatrick, D., *Evaluating training programs: the four levels*. San Francisco: Berrett-Koehler, 1994.
- Kluger, A.N. en A. DeNisi, 'The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory'. *Psychological Bulletin*, 119 (2), pp. 254-284, 1996.
- Landy, F. en J. Farr, *The measurement of work performance: methods, theory, and applications*. Orlando: Academic Press, 1983.
- London, M. en R. Beatty, '360-degree feedback as a competitive advantage'. *Human Resource Management*, 32 (2/3), pp. 353-372, 1993.
- McCarthy, A.M. en T.N. Garavan, 'Developing self-awareness in the managerial career development process: the value of 360-degree feedback and the MBTI'. *Journal of European Industrial Training*, 23 (9), pp. 437-445, 1999.
- McEnery, J. en P. Blanchard, 'Validity of multiple ratings of business student performance in a management simulation'. *Human Resource Development Quarterly*, 10 (2), pp. 155-172, 1999.
- McLean, G., M. Sytsma en K. Kerwin-Ryberg, *Using 360-degree feedback to evaluate management development: new data, new insights*. Academy of Human Resource Development, conference proceedings, 1995.
- Mulder, M., *Competentieontwikkeling in organisaties: perspectieven en praktijk*. 's-Gravenhage: Elsevier Bedrijfsinformatie, 2001.
- Murphy, K.R., J.N. Cleveland en C.J. Mohler, 'Reliability, validity, and meaningfulness of multisource ratings'. In: D.W. Bracken, C.W. Timmreck en A.H. Church (eds.), *The handbook of multisource feedback*. San Francisco: Jossey-Bass, 2000.
- Robinson, D. en J. Robinson, *Training for impact: how to link training to business needs and measure the results*. San Francisco: Jossey-Bass, 1989.
- Rosti, R. en F. Shipper, 'A study of the impact of training in a management development program based on 360 feedback'. *Journal of Managerial Psychology*, 13 (1/2), pp. 77-89, 1998.
- Sandick, A. van en A. Schaap-Neuteboom, *Rendement van een bedrijfsopleiding: een instrument voor het bepalen van het financiële rendement van trainingen*. Proefschrift, Universiteit Twente, 1993.
- Swanson, R.A. en E.F. Holton (eds.), *Human Resource Development Research Handbook: linking research and practice*. San Francisco: Berrett-Koehler Publishers, 1997.
- Velsor, E. van, J. Leslie en J. Fleenor, *Choosing 360: a guide to evaluating multi-rater feedback instruments for management development*. Greensboro, NC: Center for creative leadership, 1997.
- Yammarino, F. en L. Atwater, 'Do managers see themselves as others see them?' *Organizational Dynamics*, 25 (4), pp. 35-44, 1997.
- Yukl, G. en R. Lepsinger, 'How to get the most out of 360° feedback'. *Training*, 32 (12), pp. 45-50, 1995.



TOT SLOT

TERUGKIJKEN OF VOORUITBLIKKEN?

ANS GROTENDORST, HARM TILLEMA EN HANS DEKKER

Na zo'n exposé van nieuwe kennis, boeiende ervaringen en hoopgevende perspectieven past het nauwelijks om een epiloog te schrijven. De ondertitel van dit 'Tot slot' drukt als eerste dat dilemma uit: willen we hier terugblikken op de bijdragen van dit *HRD Thema* of willen we juist vooruitblikken, lijnen doortrekken naar de toekomst. Wat hebben we op deze plaats nog toe te voegen? En tegelijkertijd geeft deze ondertitel de kern van de bijdragen van dit themanummer weer: assessment als middel om terugblikkend te oordelen of juist als methodiek om vooruitblikkend aan ontwikkeling te werken. Na enig 'passen en meten' kwamen we tot een eensluidend oordeel: laten we besluiten met een vraaggesprek.

Hans: *Wat is assessment eigenlijk; wat is het anders dan beoordelen? Vanwaar de belangstelling ervoor? Is het de zoveelste hype?*

Harm, Ans: Assessment is een wat ongelukkig gekozen woord waarvoor we nog geen goede Nederlandse term hebben. Door een Engelse term te kiezen bouwen we kennelijk enige diffuusheid in omdat we nog niet in staat zijn het begrip adequaat en afdoende te benoemen. Langzamerhand is door het gebruik ervan en de discussie erover enige consensus ontstaan. Ons lijkt dat assessment in ieder geval twee zielen herbergt

(Tillema geef dat ook in zijn bijdrage aan). De ene ziel staat voor meting, vastlegging van prestaties of neutraler geformuleerd: van gedrag. De andere ziel is het waarderen en wegen van de prestaties in termen van standaarden en normering, zoals vastgelegd in kwalificatiestructuren of competenties. Bij elkaar genomen zou men kunnen spreken van 'gedragsevaluatie', maar die term oogt wat saai en oubollig. Veel vaker hoor je spreken over 'competentiebeoordeling' en dat klinkt al wat beter. Terecht wordt bij het gebruik van de laatste term opgemerkt: 'Wat is dan het verschil tussen assessment en beoordelen?' *Precies* beschouwd is er wel degelijk een verschil, want assessment omvat ook het meten. *Rekkelijker* bezien zijn de verschillen relatief, het oogmerk van assessment is immers steeds beoordelen.

De vaart waarmee de term 'assessment' de discussie, vooral in het reguliere onderwijs, heeft weten te veroveren, doet vermoeden dat we te maken hebben met de zoveelste hype. En die zal wel weer overwaaien, zo verwachten de preciezen. Toch denken we dat 'assessment is here to stay'. Die overtuiging vloeit voort uit onze observaties van ontwikkelingen in (het denken over) onderwijs, vooral ingegeven door externe invloeden:

- het beleid rondom een leven lang leren;
- de betekenis van de kenniseconomie;

- de noodzaak tot de ontwikkeling van praktijkrelevantere beroepsopleidingen;
- vergroting en flexibilisering van de instroom van studenten in het onderwijs;
- erkenning van reeds verworven kwalificaties;
- verantwoording van opleidingstrajecten naar de buitenwereld.

Niet alleen in het onderwijs, maar ook in arbeidsorganisaties zien we dat de aandacht voor Human Resource Development leidt tot een cruciale rol voor assessment. Daarbij gaat het om het stimuleren van loopbanen ‘in company’, de ontwikkelingen rond kenniswerk en de betekenis van coaching en motivering van medewerkers. Hoe verhoudt het feitelijk functioneren zich tot het (door de organisatie én de medewerker) gewenste functioneren? Assessment helpt die vraag te beantwoorden. Assessment speelt daardoor een steeds belangrijker rol in beslissingen over selectie, plaatsing, ontwikkeling en beloning en manifesteert zich daarmee steeds meer als een ‘blijvertje’.

Hans: *De belangstelling voor assessment leeft zowel in het bedrijfsleven als in het onderwijs. Wat zijn de overeenkomsten en verschillen in de benaderingen en de toepassingen die men kiest?*

Ans: Assessment vindt zijn oorsprong in het bedrijfsleven, in de beoordeling van de geschiktheid van kandidaten voor managementfuncties. *Byham* geeft in zijn bijdrage een mooi overzicht van de ontwikkeling van het gebruik van assessment. De nadruk ligt op voorspelling: hoeveel kans heeft iemand op een succesvolle beroepsuitoefening in bedrijf X? Daarin heeft assessment een onmiskenbare selectieve functie. Het accent is in de loop der tijd meer komen te liggen op competentieontwikkeling. *Byham* noemt de toepassing van assessment vanaf midden jaren tachtig van de vorige eeuw in het kader van ‘empowerment’ van personeel. In de vormen waarin het accent voornamelijk op ontwikkeling is komen te liggen, spreken we inmiddels over ‘development centers’.

In assessments die in de context van bedrijven plaatsvinden, is vooral belangstelling merkbaar voor de combinatie vaardigheden-eigenschappen (voorbeelden: stressbestendigheid, organisatiesensitiviteit, besluitvaardigheid, resultaatgerichtheid, analytisch vermogen, communicatieve vaardigheden). Wanneer je zo’n competentieoverzicht leest, weet je aanvankelijk niet of je te doen hebt met een manager in een bierbrouwerij, een ontwerper in een ICT-bedrijf, een leraar of een fiscaal onderzoeker. Het lijkt een truc, een kunstje, een slimme broodwinning voor organisatieadviseurs. Met een ‘toolkit’ in de hand komt men door het ganse land. De werkelijkheid is gelukkig genuanceerder, *Van der Horst* is als organisatieadviseur de eerste

om ons daarvan te overtuigen. Competenties zijn contextueel, situatiegebonden en de beoordeling ervan vergt maatwerk: in welke kenmerkende beroepssituaties moet iemand de betreffende competenties kunnen demonstreren?

Assessment is in het onderwijs in de belangstelling gekomen toen de verandering van leerstof- en vakgerichte curricula naar competentie- en praktijkgerichte curricula doorzette, zo in de jaren negentig van de vorige eeuw. Ook in het beroepsonderwijs gaat het om de vraag: in welke kenmerkende beroepssituaties moet iemand de betreffende competenties kunnen demonstreren?

Deze vraag is aan de orde in de bijdragen van *Elshout-Mohr*, *Oostdam* en *Overmaat*, van *Grotendorst* en *Schenning*, en van *Bolhuis* en *Van Royen*. Tezamen vormen deze bijdragen een kleine ‘state of the art’ in het denken over examineren en beoordelen (deugdelijk assessment) in het middelbaar en hoger beroepsonderwijs en in brancheopleidingen.

Een speciale belangstelling voor assessment ontstond toen vraagstukken op de arbeidsmarkt daartoe noodzaakten. Veelal ging het om een nijpend gebrek aan personeel, waarin het gevestigde onderwijssysteem op korte termijn niet kon voorzien. Deze urgentie heeft onder meer als katalysator gewerkt bij de ontwikkeling van EVC-procedures: het erkennen van verworven competenties, in formele opleidingen, informele leertrajecten of door werk- en levenservaring. Op grond van een ‘deugdelijk assessment’ is daardoor opeens de weg naar het beroep van leraar, verzorgende of politiemedewerker korter en aantrekkelijker geworden.

Vergeleken met het bedrijfsleven ligt in het (beroeps)onderwijs in assessments meer nadruk op ‘terugkijken’ (wat kan ik al, welke bekwaamheden heb ik?) en op de combinatie vaardigheden-kennis. Eigenschappen spelen (nog?) een ondergeschikte rol. Er is echter een groeiende belangstelling waarneembaar voor ‘leervermogen’ en de voorspellende waarde van toetsen en examens.



Harm Tillema



Ans Grotendorst

Harm: Het is opvallend dat de toepassing van assessments in het bedrijfsleven en in het onderwijs zulke onderscheiden wegen heeft genomen. Grosso modo kun je stellen dat het bedrijfsleven assessment omarmt (voorzover daar al sprake van is) voor personeelsbeoordeling, vooral in relatie tot functioneringsgesprekken. Het instrumentarium is daar ook op afgestemd, gelet op bijvoorbeeld de populariteit van het 360°-feedbackinstrument. Dit instrument is vooral een managementtool dat (gevarieerde) input levert voor een beoordelingsgesprek (zie de bijdrage van *Jellema*). De gespecialiseerde adviesbureaus richten zich vooral op selectiedoeleinden, met daarbij een grote populariteit voor het assessment center (zie de bijdrage van *Lievens* en *Kolk*). Heel anders ligt de belangstelling in het onderwijs, waar assessment vooral dient voor ontwikkeling en begeleiding (althans als we het hebben over assessment binnen onderwijsprogramma's). We zien dat in deze context het portfolio-instrument ongemeen populair is. Verder wordt assessment in het onderwijs vooral gebruikt voor verwijzing en doorstroom naar geschikte leertrajecten, onder andere bedoeld voor flexibilisering van leerwegen. Er zijn dus in dit opzicht grote verschillen tussen bedrijfsleven en onderwijs in de aard van de discussie rond assessment en de keuze of populariteit van instrumenten. De verschillen zijn te verklaren uit de functie die assessment vervult (zie de bijdrage van *Tillema* hierover) en daar is natuurlijk niets mis mee. Assessment krijgt betekenis door de wijze waarop het wordt ingezet.

Ans: Een en ander leert dat het steeds goed is te bedenken, wat een schrijver bedoelt met 'assessment'. Betekenis en reikwijdte van het begrip verschillen aanzienlijk. *Van der Horst* heeft het vooral over assessment in relatie tot beoordelen bij selectie in bedrijven. *Elshout-Mohr*, *Oostdam* en *Overmaat* gaan in op competentiegericht en formatief evalueren in het hoger beroepsonderwijs. *Jellema* bespreekt evaluatie-uitkomsten van training voor loopbanen in arbeidsorganisaties, terwijl *Grotendorst* en *Schennings* ingaan op de rol van assess-

ment in het reguliere onderwijs bij certificering en diplomering.

Hans: Wat moeten we verstaan onder 'deugdelijkheid' van assessments? Waar komt die gedachte vandaan en wie zijn betrokken bij het vaststellen van deugdelijkheid?

Harm, Ans: Deugdelijkheid verwijst naar grondigheid en degelijkheid. Daar draait het ook precies om in assessments. Laten we wel zijn, er zijn flinke en soms verstrekende beslissingen gemoeid met de uitkomst van een assessment. Op zijn minst mag je dus verwachten dat een assessment deugt voor hetgeen waarvoor het is bedoeld. Bij gedragsevaluatie en competentiebeoordeling vragen we van mensen om veel van zichzelf te laten zien, het vraagt veel tijd en inzet. Ook van beoordelaars vergen we nogal wat: nauwgezet kijken en zorgvuldig afwegen van wat zij zien aan gedrag. En van managers verlangen we dat zij de rijkdom aan informatie over functioneren naar waarde zullen schatten. Al met al zijn heel wat actoren betrokken in het assessmentproces. De vraag naar deugdelijkheid wil zeggen: er is aandacht nodig voor de kwaliteit van assessments. Dit kan naar verschillende kanten uitgewerkt zijn: is alles wat rond het assessment gebeurt helder en duidelijk? Werken de procedures, instrumenten en methoden naar behoren? Geven de uitkomsten een getrouw beeld van kwaliteiten? In de bijdrage van *Tillema* zijn criteria en kenmerken uitgewerkt voor het vaststellen van een deugdelijke assessmentpraktijk. Ook in de bijdrage van *Lievens* en *Kolk* is nadrukkelijk aan de orde hoe assessments beoordeeld kunnen worden op kwaliteit. De bijdrage van *Grotendorst* en *Schennings* gaat in op een aantal dilemma's die we tegenkomen bij het inzetten van assessments. Al deze inzichten laten onverlet dat het uiteindelijk de gebruiker is (de klant, de beoordeelde) die moet kunnen vaststellen of de persoonlijke kwaliteiten voldoende uit de verf komen. Deugdelijkheid vertaalt zich uiteindelijk in 'consumentenvertrouwen'.

Hans: Zijn we met de huidige aandacht voor assessment op het goede spoor? Komen de essentiële zaken voldoende voor het voetlicht? Vergeten we wellicht belangrijke punten?

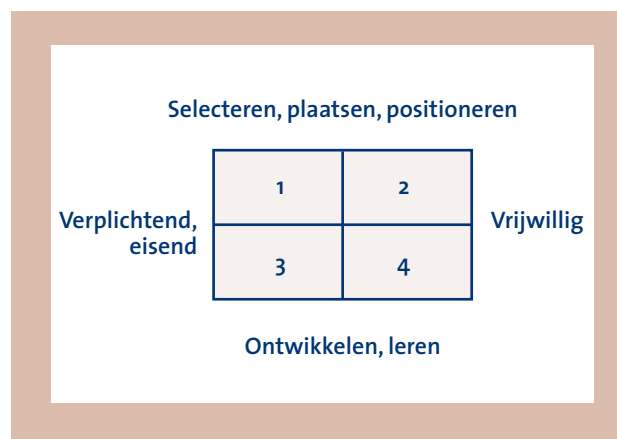
Harm, Ans: Assessment is een relatief nieuw topic in de discussies over beoordelen en toetsen. Het is een andere kijk op en manier van meten en beoordelen, zoveel is duidelijk. De afgelopen jaren zijn er dan ook veel conferenties, studiedagen en workshops gewijd aan de vraag wat assessment nu precies is, vooral in relatie tot competentiegericht opleiden en ontwikkelen. In deze adoptiefase gaat bijna als vanzelf de primaire aandacht uit naar begripsverheldering. Snel daarna gaat men over naar de hoe-vraag: Hoe doen we assessments? Welke instrumenten zijn ervoor beschik-

baar? We zien allerlei voorstellen en ontwerpen opduiken zoals het portfolio, het development center of proeven van bekwaamheid. Deze aandacht voor instrumentatie is goed te verklaren. Je wilt weten hoe iets werkt en wat het allemaal concreet voorstelt. Nu we een aantal jaren verder zijn, duiken steeds vaker vragen op als: Waartoe dient het allemaal? Wat is het nut ervan? Hoe past het bij een aantal andere wezenlijke zaken als examinering, certificering, leer- of ontwikkelingstrajecten? Welke bekwaamheden hebben assessoren nodig? De wens om duidelijkheid te verkrijgen over inpasbaarheid of integratie van assessment in een (onderwijs)organisatie komt hierin pregnant naar voren. De vraag is dan ook hoe we assessment weten in te passen in institutionele kaders. Los van de directe instrumentele vraag zal er aandacht moeten zijn voor de inbedding in een groter functioneel geheel – de verbanden met de context waarin assessment functioneert moeten helder zijn. We doelen hier op de relaties met opleiden en leren, met functioneren en loopbanen van mensen, met het beleid van de organisatie als geheel. We zullen het blikveld rondom assessment dienen te verruimen en naar de institutionele functie van assessment moeten kijken: Wie heeft er wat aan? Waarvoor dient het? Dit zijn wezenlijke vragen voor assessment. Terugblikkend op de bijdragen in de bundel zien we het instrumentele perspectief duidelijk aanwezig, met aandacht voor de werking van assessments. Toch is er ook aandacht voor de institutionele kant. De bijdrage van *Grotendorst* en *Schening* over de functie van assessment voor examinering en dilemma's daarbij, de betekenis van kwaliteitsborging bij *Tillema* en de succesvolle inzet van assessment in organisaties bij *Van der Horst*.

Waarom is de positionering van assessment in een bredere context zo belangrijk? We zien vaak bij veel vernieuwende ideeën dat ze na verloop van tijd ingekapseld raken en verworden tot een aspect en een geïsoleerd leven gaan leiden. Geleidelijk verliezen ze dan hun innovatieve waarde. Total Quality Management is zo'n voorbeeld: de totaalimpact is er inmiddels wel van af. Assessment als aspect doet tekort aan het veranderend vermogen dat er nu nog in besloten ligt, namelijk een andere kijk op ontwikkelen en beoordelen van competenties.

Hans: *Nog even naar de kwalificerende functie die assessment ook kan hebben. In de bijdragen die over beroepsgerichte leertrajecten gaan, komt die functie naar voren.*

Ans: Het is interessant om te zien hoe opleidingen er steeds beter in slagen om nieuwe vormen van beoordelen en examineren te ontwikkelen. Het gaat daarbij niet alleen maar om interessante varianten van kennisreproductie. In de nieuwe beoordelingsvormen zien we een sterke verbinding van werken en leren, school



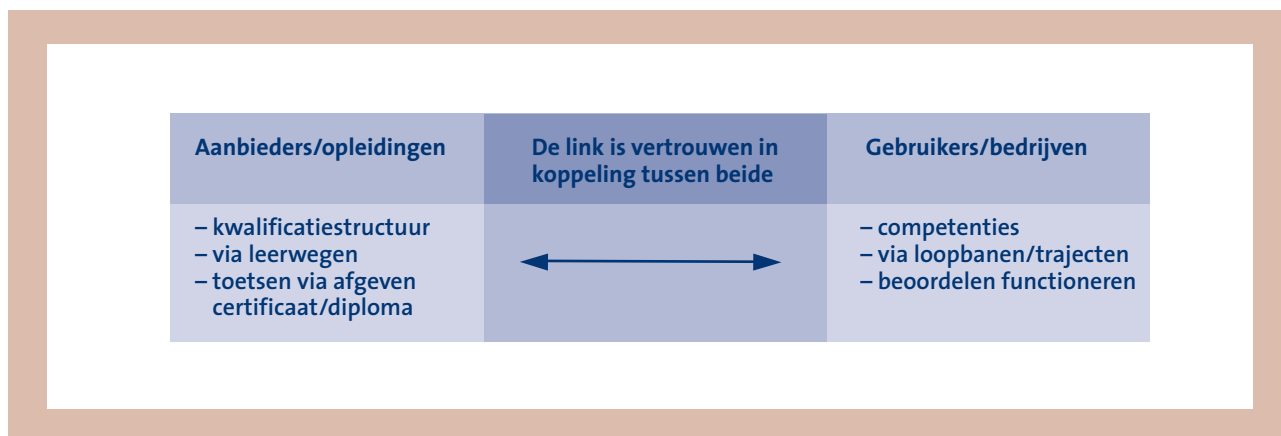
Tabel 1. Vier functies van beoordeling

en praktijk, kennis en vaardigheden, onderwijs en maatschappij. In de nieuwe beoordelingsvormen komt een heel nieuwe positionering van het onderwijs tot uitdrukking, die veelbelovend is in het licht van de opkomende kenniseconomie. De bijdrage van *Bolhuis* en *Van Royen* is een mooi voorbeeld. In een van origine gesloten en hiërarchische organisatie als de politie, lukt het om vitale maatschappelijke vraagstukken als leefbaarheid en veiligheid te verbinden aan vernieuwende concepten op het gebied van opleiden en beoordelen. De nieuwe beoordelingsvormen vormen een impuls voor het organiseren van het leren op de werkplek. Verbinding, dat is hier eigenlijk waar het om gaat.

Hans: *Assessment kent een ontwikkelingsperspectief, maar het gaat toch ook vaak om beoordelen? Hoe verhouden deze perspectieven zich tot elkaar?*

Harm: Er treedt inderdaad vaak verwarring op over de verschillende functies die assessment kan hebben. Wellicht kan een tabel dit het beste verduidelijken. Het gaat in feite steeds om verschillende benaderingen of manieren van kijken naar assessment. Aan de ene kant valt er een onderscheid te maken tussen ontwikkelend of toetsend beoordelen. Aan de andere kant kan het gaan om het opgelegde, verplichtende karakter dat eigen is aan programma- of certificeringseisen, de zelfbeoordeling en feedbackverzameling over het eigen functioneren. Dit is schematisch weergegeven in tabel 1.

Het is goed zich steeds bewust te zijn van het kwadrant waarin men zich bevindt. In kwadrant 1 heeft men te maken met bijvoorbeeld het dossier of de EVC-procedure waarin men ten behoeve van een instantie moet bewijzen of men voldoet aan selectie-eisen. In kwadrant 3 gaat het bijvoorbeeld om een eindtoets die men aflegt om te laten zien of men de cursuseisen heeft behaald. In kwadrant 2 gaat het om een CV-achtig verslag of proeve van bekwaamheid, waarin men



Tabel 2. Verschillende verantwoordelijkheden rond assessment

zelf aantoonbaar welk niveau men heeft bereikt. In kwadrant 4 ten slotte gaat het om persoonlijke ontwikkelingsplannen en portfolio's waarin men bijhoudt hoe men zich ontwikkeld heeft in een bepaalde periode.

Hans: Bij assessment gaat het om het 'spiegelen' van iemands competenties aan een standaard, een norm, aan beoordelingscriteria. De vraag is dan natuurlijk: Wie bepaalt de standaard? Wie is normstellend?

Ans: Dat is lang niet altijd duidelijk. Een gekozen beoordelingsnorm is idealiter de uitkomst van een zorgvuldige, gezamenlijke analyse met betrokkenen: managers, medewerkers, specialisten, ontwikkelaars. In die analyse komen vragen aan de orde als: Welke kenmerkende situaties doen zich voor in de praktijk? Welke ontwikkelingen zijn daarin te verwachten? Welke opgaven en problemen moeten medewerkers, managers, professionals tot een oplossing zien te brengen? Op welke competenties wordt daarbij een beroep gedaan? Hoe kunnen de medewerkers, managers, (aankomend) beroepsbeoefenaars deze competenties demonstreren en aan welke eisen moeten zij daarbij voldoen?

Het doorlopen van zo'n procedure vraagt tijd en

eigenlijk loop je altijd een beetje achter de feiten aan. Het zou dus bij voorkeur een permanent proces moeten zijn. Naarmate het domein waarover je uitspraken moet doen groter is (zoals in het beroepsonderwijs), wordt de beschrijving van beoordelingscriteria waarin iedereen zich kan vinden lastiger, zo niet onmogelijk. Op z'n best kun je het eens worden over beschrijvingen van te verwerven competenties. De inhoud van de beoordeling moet je niet willen standaardiseren, competenties en beoordelingsnormen zijn situationeel en daarbij is inbreng en betrokkenheid van lokale actoren gewenst. Wat je wel kunt standaardiseren zijn kwaliteitsnormen, processen en procedures.

Harm: Dit is een wezenlijke maar ook ingewikkelde vraag, vooral omdat meerdere partijen en meerdere verantwoordelijkheden een rol spelen. Het gaat eigenlijk om de correspondentie tussen wat de aanbieder levert en wat de vrager wil met betrekking tot competentieverwerving en certificering van verworven kwalificaties. Aan de ene kant hebben we de aanbieder, de opleidingen in feite, die meestal via opgestelde kwalificatiestructuren (voor een branche) een leerweg programmeren en daarop certificaten afgeven. Aan de andere kant gaat het om de gebruikers en afnemers in bedrijven die beschikbare competenties of eerder verworven kwalificaties goed zichtbaar en redelijk uniform en breed getoetst willen zien. In tabel 2 is dat schematisch weergegeven.

We zien dat de link of correspondentie tussen beide op dit moment nog lang niet in voldoende mate erkend of in vertrouwen geaccepteerd wordt. Enerzijds hebben we de aanbieders die op een lewegafhankelijke manier toetsen welke (aangeboden) kennis en vaardigheden beheerst worden en anderzijds de afnemers die op eigen gronden naar de functionaliteit van de verworven competenties kijken. Het punt is natuurlijk hoe de link of de correspondentie tot stand wordt gebracht. Een manier zou zijn dat door de aanbieders



Hans Dekker

voldoende vertrouwen wordt geboden in het assessment van de via de opleiding verworven competenties. Dus, opleidingen die oog hebben voor de standaarden (lees: eisen) die de afnemers stellen, houden in het assessment van hun programma's zodanig rekening met deze eisen dat er vertrouwen ontstaat in de uitslag van hun toetsing. Kortom, het antwoord op de vraag is eigenlijk: in de afstemming ligt de bepaling van de norm.





LITERATUUR- EN WEBWIJZER

ANS GROTENDORST, HARM TILLEMA EN HANS DEKKER

Ter afsluiting van deze *HRD Thema* willen we lezers die zich verder willen verdiepen in het vraagstuk van 'deugdelijk assessment in arbeidsorganisaties en onderwijs' suggesties doen voor die verdieping. De onderstaande bronnen zijn bijeengebracht door de auteurs van dit katern. De bronnen zijn geordend naar enkele subthema's, te weten:

- *Competentieontwikkeling en beoordelen*. Hier komen bijdragen aan bod die het beoordelingsthema breed benaderen.
- *Relatie assessment en beoordeling*. Bijdragen in dit cluster gaan met name in op de kwaliteitsvraag: aan welke criteria moet deugdelijke beoordeling voldoen, hoe kom je tot een valide beoordeling?
- *Ontwikkelingsgericht beoordelen in development centers*. Deze categorie bijdragen belicht het development center als specifieke vorm van beoordeling: voorstellen, eisen, vormen, voorbeelden.
- *Ontwikkelingsgericht beoordelen via portfolio's*. Hier wordt het portfolio nader belicht, op dezelfde wijze als het development center.

COMPETENTIEONTWIKKELING EN BEOORDELEN (ALGEMEEN)

- *Competentiegericht leren en werken*, Utrecht: Lemma.

Deze reeks bestaat uit drie bundels waarin uitgebreid aandacht wordt besteed aan assessment en beoordelen in met name het hoger onderwijs en de beroepspraktijk. Deze drie bundels komen voort uit een reeks conferenties over competentiegericht leren en werken in het hbo en bieden tal van voorbeelden en uitwerkingen rond competentiegerichte vormen van onderwijs en beoordeling. Elke bundel bevat enkele artikelen over invoering van assessmentinstrumenten. De bundels zijn:

- Buskermolen, F., B. de la Parra en R. Slotman, *Het belang van competenties in organisaties*. Utrecht: Lemma, 1999.
- Schlusmans, K., R. Slotman, C. Nagtegaal en G. Kinkhorst, *Competentiegerichte leeromgevingen*. Utrecht: Lemma, 1999.
- De la Parra, B., R. Slotman, T. Spannenburg en H.H. Tillemma, *Managen van competenties in organisaties*. Utrecht: Lemma, 2000.

- Elshout-Mohr, M. en R. Oostdam, *Assessment van competenties in een dynamisch curriculum*. Amsterdam: SCO-Kohnstamm Instituut, 2001.

Dit rapport (147 pagina's) bundelt uiteenlopende inzichten over beoordelen en assessment binnen het hoger onderwijs en nodigt sterk uit tot discussie. Aan de orde komen drie toetsmodellen die elk hun eigen toepassingsgebied hebben. Het eerste model (het één-op-één model) betreft het toetsen van de mate van beheersing van discrete leerdoelen, zoals de beheersing van vakspecifieke kennis en vaardigheden. Het tweede model (het er-leiden-vele-wegen-naar-Rome model) betreft beoordeling van de beheersing van generieke leerdoelen, zoals probleemoplossings- en communicatievaardigheden. Het derde model (het Rome-en-Rome-is-twee model) betreft assessment van gesitueerde leerdoelen, zoals beroepscompetenties. Van elk model wordt de achterliggende beoordelingsgedachte besproken, het beoordelingsscenario, de gebruikte instrumenten en de kwaliteitseisen. Ook wordt ingegaan op praktische problemen en nieuwe trends. Het derde model is in het rapport het meest uitgewerkt op grond van praktijkervaringen in de Educatieve Faculteit Amsterdam (EFA).

Het rapport is vooral bedoeld als bron van informatie en ideeën voor docenten en stafmedewerkers van instellingen voor hoger beroepsonderwijs, die een overstap willen maken naar competentiegericht opleiden en tot het inzicht zijn gekomen dat het organiseren van passende beoordelings- en assessmentprocedures een speerpunt kan vormen in hun beleid.

Het rapport is te bestellen bij SCO-Kohnstamm Instituut, Wibautstraat 4, 1091 GM Amsterdam, tel. 020-5251201 (onder vermelding van rapportnummer 605).

- Klarus, R., *Competenties erkennen*. 's-Hertogenbosch: Cinop, 1998.

Dit is een overzichtswerk van de hantering van beroepsprofielen en erkenning van kwalificaties in het beroepsonderwijs. Klarus besteedt ruim aandacht aan de totstandkoming (selectie en verzameling) van competentieprofielen.

- Klarus, R., H.H. Tillema en J. Veenstra, 'Beoordelen: met competentieprofielen of kwalificatiestructuren'. *Opleiding en Ontwikkeling* 12 (11), pp. 15-20, 1999.

In dit artikel doen de auteurs verslag van empirisch onderzoek naar de overeenkomsten en verschillen in de aanpakken die in het bedrijfsleven en in het onderwijs worden gehanteerd om beroepsprofielen op te stellen.

- Nedermeijer, J. (red.), *Beroepscompetenties en academische vorming in het hoger onderwijs*. Wolters-Noordhoff: Groningen, 2000.

Dit is een inleiding over de betekenis van beroepscompetenties voor onderwijs en de manieren van verzameling van competentieprofielen voor de opbouw van het curriculum.

RELATIE ASSESSMENT EN BEOORDELING

- Dochy, F.J.R.C. en T.R. de Rijke (red.), *Assessment centers in de praktijk: nieuwe toepassingen in opleidingen, onderwijs en HRM*. Utrecht: Lemma, 1995.

Dit overzicht van nieuwe vormen van toetsing en beoordeling, met toelichtingen over hun achtergrond en voor- en nadelen alsmede enkele onderzoeksgegevens, is geschikt voor gebruik in het onderwijs.

- Dochy, F.J.R.C., L. Heylen en H. van de Mosselaer (red.), *Assessment in onderwijs, nieuwe toetsvormen en examinering in studentgericht onderwijs en competentiegericht onderwijs*. Utrecht: Lemma, 2002.

Dit boek geeft een overzicht van de nieuwste ontwikkelingen op het gebied van assessment. Ook worden de consequenties voor het huidige onderwijs in kaart gebracht. Er wordt aandacht geschonken aan de kwaliteit van nieuwe assessmentvormen en aan de criteria die hiervoor gewenst zijn. De auteurs bieden een overzicht van diverse assessmentvormen, geven richtlijnen voor het gebruik hiervan en illustreren dit met praktische voorbeelden.

- Hofstee, W.K.B., *Principes van beoordelen*. Lisse: Swets & Zeitlinger, 1999.

In dit boek wordt de problematiek van assessmentprocedures besproken vanuit een brede kennis en ervaring met diverse toetspraktijken. Prof. dr. W.K.B. Hofstee is hoogleraar in de psychologie aan de Rijksuniversiteit Groningen. Het thema 'beoordeling' is een rode draad in zijn werk op het gebied van persoonlijkheids-, onderwijs- en selectiepsychologie en methodologie. Het boek is geschreven op basis van een speurtocht naar gemeenschappelijke principes die gelden voor beoordelingen in uiteenlopende contexten, zoals het beoordelen van sollicitanten en werknemers door arbeidsorganisaties, van studenten en leerlingen door onderwijsinstellingen en subsidieaanvragen door overheden. Veel aandacht wordt besteed aan het onderscheiden van de verschillende actoren in assessments en de methodische en ethische grondslagen van

hun handelen. Ook gaat Hofstee in op de constructie, verwerking en rapportage van vragenlijsten, op beoordeling via experimenten, testen, toetsen en normen en op cesuren en standaarden.

Het boek is geschikt voor iedereen die verantwoordelijkheid draagt voor selectie- en beoordelingsprocedures en voor studenten die zich hierop willen voorbereiden.

- Straetmans, G.J.J.M. en P.F. Sanders, *Beoordelen van competenties van docenten*. Den Haag: Programmamanagement EPS/HBO-raad, 2001.

Deze publicatie richt zich op assessmentinstrumenten voor docenten, maar is ook toepasbaar op andere beroepen. De auteurs gaan in op verschillende functies van en methoden voor competentiebeoordeling en de kwaliteitseisen die aan deze methoden kunnen worden gesteld. De boodschap is dat er geen sprake is van één 'beste' methode, maar dat het de voorkeur heeft om de zogenoemde methodenmix toe te passen, om alle zwakke kanten van de ene methode te compenseren met sterke kanten van een andere methode. De publicatie eindigt met een aantal beschrijvingen en evaluaties van gebruikte competentie-instrumenten van docenten.

Deze publicatie is te vinden op en te downloaden van de EPS-website www.educatiefpartnerschap.nl.

- Tillema, H.H., *Assessment van competenties; van beoordelen naar ontwikkelen*. Deventer: Kluwer, 2001.

Dit boek biedt een overzicht van het gebruik van assessment in organisaties en plaatst dit met name in het perspectief van Human Resource Development van organisaties. Het gaat in op de beoordeling van assessment door middel van een geschiktheidsbeoordelingstest.

- Vleuten, C.P.M. van der en E.W. Driessen, *Toetsing in probleemgestuurd onderwijs*. Groningen: Wolters-Noordhoff, 2000.

Met probleemgestuurd onderwijs worden specifieke onderwijsdoelen beoogd. Wil deze vorm van onderwijs slagen, dan moet het toetsprogramma de onderwijsdoelen respecteren, of nog liever, bevorderen. Dat vereist een andere dan de traditionele aanpak van toetsing. Dit boek bevat een overzicht van de wijze waarop toetsing in probleemgestuurd onderwijs kan worden aangepakt. Het bevat een overzicht van de huidige inzichten uit wetenschappelijk onderzoek en ontwikkeling op het gebied van nieuwe toetsvormen, een aantal praktische voorbeelden uit verschillende disciplines en een aantal algemene wenken voor de opzet van een dergelijk toetsprogramma.

ONTWIKKELINGSGERICHT BEOORDELEN IN DEVELOPMENT CENTERS

- Boot, C. en H.H. Tillema, *Competentiegericht beoordelen in hoger beroepsonderwijs*. Utrecht: Lemma, 2001.

Dit boek gaat in op het development center als ontwikkelings- en begeleidingsinstrument. Tevens benoemen de auteurs succesfactoren die bij de ontwikkeling van een dergelijk systeem aandacht behoeven: de rol van het management en docenten, de flexibiliteit van de organisatie en de overtuiging bij de student van het nut van beschreven instrumenten. Er komen zowel ontwikkelaars en gebruikers aan het woord.

- Eringa, K., J. Rietveld en W. Zwaal, *Assessment en development centers voor studenten*. Groningen: Wolters-Noordhoff, 2000.

Studenten moeten aantonen dat ze beschikken over de competenties om adequaat te kunnen handelen in bepaalde kritische werksituaties. Een assessment en development center is bij uitstek de manier om competenties te toetsen. Dit boek biedt praktische aanbevelingen voor het ontwikkelen, uitvoeren en evalueren van assessment en development centers, die van toepassing kunnen zijn bij allerlei beslissingen in het studietraject: selectie, erkenning van eerder verworven competenties, studieloopbaanbegeleiding, flexibele studieroutes en vaardigheidsbeoordeling. Ook vragen rondom kosten, haalbaarheid, organisatie en benodigde faciliteiten komen uitvoerig aan de orde.

- Jansen, P.W.G. en F. de Jongh, *Assessment centers, een open boek*. Baarn: Spectrum, 1993.

Dit is een inleidende tekst over de verschillende gebruiksmogelijkheden van het assessment center, met bijzondere aandacht voor de wijze van afname en het verloop van de procedure.

- Tillema, H.H., 'Assessment / development centers – sturen op prestaties'. *Gids voor de Opleidingspraktijk* 28, (1.56), pp. 1-24, 1999.

Een inleidende tekst over de kenmerken en mogelijkheden van assessment centers dan wel development centers voor het beoordelen van leerresultaten en leervorderingen, vooral gericht op de beroepspraktijk.

- Tillema, H.H. (red.), *Development centers, competentie-ontwikkeling in organisaties*. Deventer: Kluwer Bedrijfswetenschappen, 1996.

Overzicht van verschillende vormen van beoordeling van competenties. Enkele instrumenten worden besproken, alsmede hun inzet en gebruik voor het ontwikkelingsgericht vaststellen van (leer)vorderingen.

ONTWIKKELINGSGERICHT BEOORDELEN VIA PORTFOLIO'S

- Evelien, F. en J. van Tartwijk, 'Overwegingen bij het gebruik van portfolio's binnen een universitaire lerarenopleiding'. *Tijdschrift voor lerarenopleiders* 21, pp. 38-46, 2000.

Praktijkervaringen bij de inzet van portfolio's in de opleiding van studenten, alsmede een afweging van het nut van portfolio's.

- Tillema, H.H., 'Het portfolio als assessmentinstrument voor potentieelontwikkeling'. *Checklisten voor personeelsmanagement* 3, (2.01), pp. 1-20. Deventer: Kluwer Bedrijfswetenschappen, 1997.

Een praktische uitwerking van de inzet van het portfolio-instrument voor beoordeling, waarbij onderscheid is gemaakt naar verschillende typen portfolio's.



OVER DE AUTEURS

B.G. Bolhuis is senior toetskundige bij het Centraal Examenbureau Politie van het LSOP Politie Onderwijs- en Kenniscentrum te Apeldoorn. Hij is projectleider voor alle proeven van bekwaamheid die worden ontwikkeld voor het Instituut voor de BasisPolitiefunctie (IBP). Hiervoor was hij docent aan meerdere IBP's en daarvoor praktijkdocent in een korps. Hij is nauw betrokken bij de ontwikkeling van het initiële politie-onderwijs en examinering. Hij is executief politie-ambtenaar (inspecteur van politie) en werkt nog regelmatig in een korps in de basispolitiezorg om de praktijkervaring op peil te houden.

E-mail: Bart.Bolhuis@lsop.nl

Dr. M. Elshout-Mohr is werkzaam bij het SCO-Kohnstamm Instituut en bij het Instituut voor de Leraren-opleiding (ILO) van de Universiteit van Amsterdam. Zij is afgestudeerd in de cognitieve psychologie en methodenleer aan de Universiteit van Amsterdam en promoveerde op het proefschrift *Training in denkstrategieën* (1976). Veel van haar onderzoek betreft leren en instructie in het hoger en middelbaar onderwijs, met een accent op de verwerving van hogere-ordevaardigheden en -competenties, zoals reflecteren en 'leren leren'. Ook de problematiek van het meten en beoordelen van deze vaardigheden behoort tot haar onderzoeksgebied.

E-mail: marianne@educ.uva.nl

A. Geldermans (politicoloog) is werkzaam als eindredacteur bij het *HBO-Journaal*, een vakblad voor het hoger beroepsonderwijs. Daarnaast werkt zij als freelance journalist voor verschillende bladen.

Drs. A. Grotendorst is partner van Kessels & Smit *The Learning Company* en lid van de Onderwijsraad. Daarnaast is zij hoofddocent van de Foundation for Corporate Education (FCE). Zij was voorzitter van de raadscommissie die het examenadvies heeft voorbereid.

E-mail: grotendorst@kessels-smit.nl

Drs. A.C. van der Horst is sinds 1989 werkzaam als senior consultant bij GITP International. Hij geeft leiding aan de adviesgroep GITP Zeist.

E-mail: a.c.van.der.horst@gitp.nl

Drs. F. Jellema werkt op het moment als AIO aan de faculteit der Toegepaste Onderwijskunde van de Universiteit Twente. Hiervoor heeft zij bedrijfskunde gestudeerd aan de Open Universiteit. Ze is bezig met onderzoek naar opleidingsevaluatie en meer specifiek met de vraag of 360°-feedback hiervoor een geschik-

te methode is. Zij hoopt dit onderzoek in februari 2003 af te sluiten met een promotie.

E-mail: f.a.jellema@edte.utwente.nl

Dr. N. Kolk promoveerde in 2001 op een proefschrift naar de validiteit van het assessment center. Momenteel werkt zij als consultant bij Berenschot Business Management voor de adviesgroep The Change Factory. Zij houdt zich bezig met organisatieverandering en strategie-implementatie.

E-mail: N.Kolk@Berenschot.com

Dr. F. Lievens promoveerde in 1999 op een proefschrift naar de begripsvaliditeit van assessment centers. Hij is nu universitair hoofddocent aan de Universiteit Gent (Vakgroep Personeelsbeleid en Arbeids- en Organisatiepsychologie). Zijn onderzoeksinteresses situeren zich in het domein van selectie en rekrutering.

E-mail: filip.lievens@rug.ac.be

Dr. R. Oostdam studeerde Nederlands en taalbeheersing aan de Universiteiten van Leiden en Amsterdam. Hij promoveerde op het onderzoek *Argumentatie in de peiling* (1991). Sinds 1983 is hij werkzaam als wetenschappelijk projectleider bij het SCO-Kohnstamm Instituut van de Universiteit van Amsterdam en sinds 1997 werkt hij tevens als programmaleider/account manager van de onderzoeksgroep Taalonderwijs. Hij publiceert nationaal en internationaal op het gebied van taalonderwijs, literatuuronderwijs en literaire respons, cognitieve leerprocessen, intelligentie en leerpotentieel, curriculumontwikkeling en assessment. Daarnaast is hij medeauteur van *Taaldomein*, een methode Nederlands voor het voortgezet onderwijs.

E-mail: ron@educ.uva.nl

Dr. M. Overmaat is als onderwijspsycholoog werkzaam bij het SCO-Kohnstamm Instituut. Zij promoveerde op het onderzoek *Schrijven en lezen met tekstschema's* (1996). Als projectleider is ze betrokken bij het flankerend onderzoek van het vernieuwingsproces op de Educatieve Faculteit van Amsterdam. Haar onderzoeksgebied betreft leren en instructie, de laatste jaren met nadruk op krachtige leeromgevingen en competentiegericht opleiden.

E-mail: marjan@educ.uva.nl

Drs. R.A. van Royen heeft onderwijskunde gestudeerd aan de Universiteit Utrecht, was docent bij een lerarenopleiding van de Hogeschool van Utrecht en is werkzaam als toetskundige bij het Centraal Examen-

bureau Politie van het LSOP Politie Onderwijs- en Kenniscentrum te Apeldoorn. Hij is met name betrokken bij de ontwikkeling van de proeven van bekwaamheid voor het initiële politieonderwijs voor politiekundige bachelor en politiekundige master. Tevens is hij betrokken bij de proeven van bekwaamheid voor de postinitiële leergangen.

E-mail: Richard.van.Royen@lsop.nl

Dr. S. Schenning is momenteel medewerker van de wetenschappelijke staf bij de Onderwijsraad. Zij was projectleider van het examenadvies. Hiervoor ondersteunde zij als stafmedewerker de inspectie BVE bij de ontwikkeling van het vernieuwde inspectietoezicht. Verder was zij werkzaam als docent en trainer, onder andere bij het ROC Albeda College in Rotterdam. Zij is gepromoveerd op de verwerving van het Nederlands als tweede taal door Marokkaanse en Turkse volwassenen.

E-mail: s.schenning@onderwijsraad.nl

Dr. H. Tillema is werkzaam als hoofddocent aan de Universiteit Leiden op het gebied van opleiding en ontwikkeling. Zijn specialisme ligt op het gebied van de competentieontwikkeling van mensen in organisaties, in het bijzonder waar het gaat om de vaststelling van hun kwaliteiten en talenten. Hij richt zich daarbij met name op de mogelijkheden die mensen hebben in hun organisatie om competenties verder te ontwikkelen. Hij heeft zich via onderzoek vooral beziggehouden met het ontwikkelen en invoeren van geschikte assessmentinstrumenten voor competentieontwikkeling. Vanuit zijn expertise is hij als consultant betrokken bij diverse projecten rond assessment en ontwikkeling in verschillende soorten organisaties.

E-mail: h.tillema@12move.nl



Reeds eerder verschenen titels in de reeks HRD Thema:

Van oude fundamenten naar nieuwe richtingen (0/2000)

- Door kennis te delen krijg je meer / *R.J.L. Visser*
- Wat vinden HRD-lezers belangrijk? / *Drs. S. Wagenaar, drs. S. Tjepkema en M. Drent*
- Van oude fundamenten naar nieuwe richtingen / *A. Mulders*
- Even voorstellen: HRD Thema en de redactie / *Drs. J.M. Dekker, dr. R.F. Poell, drs. S. Tjepkema, drs. S. Wagenaar*

Het ontwerpen van leertrajecten (1/2000)

Gastredactie: dr. N.M. Nieveen

- Ontwerpen van leertrajecten: Laten ontwerpen door professionals of zelf ontwerpend leren? / *Dr. N.M. Nieveen en dr. R.F. Poell*
- Ontwerpbependingen in opleidingspraktijken / *Dr. J.I.A. Vischer-Voerman*
- Het ontwerpproces als leerproces / *Prof. dr. J.W.M. Kessels*
- Balanceren tussen ruimte en structuur: Lerarenopleiders van de Educatieve Faculteit Amsterdam als curriculumontwikkelaars / *Drs. L.E. Odenthal, dr. W.A.J.M. Kuiper, dr. J.M. Voogt en drs. S.M.M. Terwindt*
- Operator 2000: De lerende als ontwerper van zijn eigen opleidingstraject / *H.K.M. Arens en A. Wegewijs*
- Ontwerpen van leerprocessen voor het creëren van nieuwe kennis / *Drs. S. Wagenaar en dr. P. Keursten*
- Het faciliteren van action learning: De rol van de leercoach / *Dr. J.A. O'Neil*
- Leertrajecten ontwerpen op velerlei wijze / *Dr. R.F. Poell en dr. N.M. Nieveen*

Stimuleren van kennisproductiviteit (2/2000)

Gastredactie: drs. A.W.M. Bruining

- Kennisproductiviteit stimuleren: Van voor naar achter, van links naar rechts / *Drs. S. Wagenaar en drs. A.W.M. Bruining*
- Principes van kennisoverdracht / *Dr. J.H. Oldenkamp*
- Complexiteit en kennisproductie, een goed huwelijk? Over ziens- en zijnswijze / *Drs. J. Peters en drs. R.A.E. Wetzels*
- Kennisintensief = kenniseffectief? / *M. Hoekstra en R. Schouten*
- Kennisproductiviteit aan de frontlinie / *Drs. A.W.M. Bruining*
- In vijf stappen naar zeven leerfuncties: Een praktisch instrument voor het bevorderen van kennisproductiviteit / *Drs. M. Talma en drs. P. Stienstra*
- Narrenkap en toverstaf. Of hoe je kennisproductiviteit narratief benadert / *Dr. J.P.M. Schrijvers*
- Kennisproductie en menselijke dynamiek. Of waarom sommige creatieve mensen niet van brainstormen houden / *F. de Mink*
- Doe wat je predikt! Maar hoe doe je dat als adviseur? / *J. van den Berg en M. Hoeffgen*
- Kennisproductiviteit in groepen / *Drs. T.M. van Engers, L. Vork en M. C. Puerta Melguizo*

Loopbaanontwikkeling (3/2000)

Gastredactie: drs. M.A.C.T. Kuijpers en dr. B.I.J.M. van der Heijden

- Loopbaanactualisatie: loopbaanontwikkeling voor de moderne werknemer? / *Drs. M.A.C.T. Kuijpers*
- Loopbaanontwikkeling: de rol van de werknemer en de organisatie / *Dr. B.I.J.M. van der Heijden*

- Zelfreflectie: makkelijker gezegd dan gedaan/*Dr. F. Meijers*
- Werkexploratie: het verkennen van mogelijkheden voor ander werk/*Dr. J. Onstenk*
- Loopbaansturing: de centrale rol van interne en externe dialoog/*Dr. R.P.J.M. van de Loo*
- Zelfprofilering: waarom en hoe?/*Dr. W. Altink en drs. R. Bos*
- Praktijkschets van een retailorganisatie: loopbaanontwikkeling binnen KBB/*C. Meinhardt*
- Tussen droom en daad: fasen in loopbaanoriëntatie/*Drs. F. Valkenburg*

Klantleren (1/2001)

Gastredactie: drs. A.J. Ardon en drs. ir. H.G. Westendorp

- Het belang van interactief klantleren ... in het perspectief van kenniseconomie en netwerksamenleving/*Drs. ir. H.G. Westendorp*
- De competenties van de klant benutten/*C.K. Prahalad en V. Ramaswamy*
- Benut u uw kansen op klantleren?/*Drs. J.M. Dekker en dr. O. Donnenberg*
- KPN creëert de toekomst NU/*Drs. B.G. Parson en drs. A.C. Boon*
- Leren (van) luisteren naar de consument/*Ir. B. Classen*
- Klantleren binnen politie, zorg en stadsontwikkeling: drie variaties op één thema/*A. Mulders*
- Met de klanten leren ... of hoe men bij het rijden nieuwe wegen kan inslaan/*J. Berger*
- Directe klantenfeedback/*Dr. M. Weiss*
- Klantleren: de kunst van het ontvangen/*Drs. A. de Jong*
- Zijn we wel klaar voor de klant? ... over de voorwaarden om klantleren mogelijk te maken/*Drs. A.J. Ardon*

Beroepsonderwijs (2/2001)

Gastredactie: dr. J. Brandsma

- Beroepsonderwijs en volwasseneneducatie: een sector in beweging/*Dr. J. Brandsma*
- De impact van een leven lang leren op de curricula in het reguliere beroepsonderwijs/*Drs. A.F. Westerhuis*
- Erkenning van Verworven Competenties (EVC) in bedrijf en school/*R. Klarus en H. Smulders*
- Krachtige leeromgevingen in het beroepsonderwijs/*Dr. E. de Bruijn*
- Assistentenopleidingen (niveau 1): vangnet of voorbereiding?/*Dr. J. Brandsma*
- De praktijk van een leerbedrijf. Een interview met de heer Jansen, ondernemer supermarkt C-1000/*Dr. J. Brandsma*
- De ontwikkeling van ROC's tot regionale kennisnetwerkers/*L. Nieuwenhuis*
- Contractonderwijs binnen ROC's. Een interview met Hans Smit van Deltion College Training en Advies/*Dr. J. Brandsma*

Leren over grenzen heen (3/2001)

Gastredactie: drs. G.M. van Dee

- Leren over grenzen heen: meer dan het overbruggen van taal- en cultuurverschillen?/*Drs. G.M. van Dee*
- Samenwerken in verbetersteams. Wat betekent het om 'Orama' (visie) te ontwikkelen op werken in teams in een Griekse brouwerij?/*Drs. R.A.M. van Noort*
- Leren over grenzen heen: omgaan met diversiteit/*Drs. W.S. Renes*

- 'U durft, u neemt het risico om een dag door te brengen met een Nederlander'. Het faciliteren van leerprocessen in een andere taal en cultuur/*Drs. W.J. Velthuis*
- Nieuwe aanpak voor multinationaal opleiden/*Mr. ing. W.A.M. Heine*
- Online leren in internationale natuurbescherming/*Drs. J. Breman*
- HRD tussen Twente en Moskou. Een voorbeeld van een internationaal opleidingstraject over Human Resources Development/*Drs. C.A. van Vilsteren*
- Online coaching in een internationale leeromgeving/*Dr. S. van de Bunt-Kokhuis*
- 'Opleiden is uitnodigen tot veranderen'. IFC-methode KPN Opleidingen overstijgt muren tussen culturen/*R.J.L. Visser*
- Leren over grenzen heen: succes bereiken met belangstelling en respect/*Drs. G.M. van Dee en drs. S. Wagenaar*

Persoonlijke ontwikkeling (4/2001)

Gastredacteur: dr. Th. van Dellen

- Voorbeschouwing op een weerbarstig thema/*Dr. Th. van Dellen*
- Persoonlijke ontwikkeling: een panacee voor HRD?/*Dr. Th. van Dellen*
- Persoonlijkheidsontwikkeling/*Prof. dr. F. De Fruyt*
- 'Een manager is een meneer de champion'. Een interview met Jan Schouten/*Dr. Th. van Dellen en R. Rovers*
- Persoonlijke ontwikkeling: een kwestie van intelligentie, persoonlijkheid en/of sociale vaardigheden?/*Drs. W. Bloemers en dr. P. van der Maesen de Sombreff*
- 'Passie is een economische noodzaak'. Een interview met Joseph Kessels/*Dr. Th. van Dellen en R. Rovers*
- De noodzaak tot een waardengedreven benadering van HRD/*Drs. E.L. Kinds*
- Organisatiebelang en individueel belang, balanceren in een dilemma/*D. Ratering en drs. L. Halbertsma*
- 'Organisaties zijn niet erg liefhebbend'. Een interview met Harry Starren/*Dr. Th. van Dellen en R. Rovers*
- Wat je doet, doet ertoe! Persoonlijke ontwikkeling bij de Nederlandse Spoorwegen/*Drs. R.A.M. van Noort, drs. K. van Ginkel, drs. S. Ringeling en ir. H. Herbert*
- Persoonsontwikkeling voor wie?/*Prof. dr. W.K.B. Hofstee*

Organiseren van de HRD-functie (1/2002)

Gastredactie: drs. P. Bührs

- De HRD-functie, wat zien we aan ontwikkeling?/*Drs. P. Bührs en drs. J.M. Dekker*
- Human Resource Development en de corporate university/*J. Walton*
- Werken en groeien met competence centers/*Drs. M. Bolhuis en dr. P. Keursten*
- Corporate virtual universities: Een vierdelig fuga van leren, kennis, persoonlijke en organisatieontwikkeling/*Prof. dr. W. Baets*
- Wie organiseren leersystemen? De onderschatte invloed van managers en werknemers/*Dr. F.J. Van der Krogt*
- Doelgericht experimenteren moet worden bevorderd. Een rondetafelbijeenkomst met NS en Heineken/*R. Breedveld en drs. P. Bührs*
- De effectieve opleider: van plannen, sturen en beheersen naar faciliteren/*Dr. J. van der Waals*

- De HRD-functie belicht, medewerker én manager aan het stuur / *Drs. P. Bührs en drs. J.M. Dekker*

Leren van professionals: beroepsgericht of organisatiegericht?

(2/2002)

Gastredactie: drs. L. Tijmens

- Leren van professionals binnen het beroep / *Dr. R. Poell en drs. L. Tijmens*
- Leren van advocaten: de invloed van tradities en trends / *Drs. L. Derksen en mr. M.D. Vreugdenhil*
- Leren in de accountancyprofession: Waar verwerven accountants hun vaktechnische en sociale competenties? / *Dr. R. Vaatstra*
- Leren diverse professionals anders? Opleiders van huisartsen, top-zwemcoaches en casinomedewerkers geïnterviewd / *Drs L. Tijmens, dr. R. Poell en drs. I. Schoemakers*
- Leren in een beroep dat niet bestaat: de casus informatici / *Dr. R. van Dael*
- Leren van organisatieadviseurs in het beroep: 'Van harde kennis tot zachte wijsheid' / *Dr. W.J.P.A. van Beers*
- Leren van docenten in het beroep: momenten van reflectie? / *Drs. H.C.M. Balvers en drs. L. Tijmens*
- Leren en de beroepspraktijk: een onderzoek naar vier beroepen / *Dr. B.J. Daley*
- Leren van professionals: beroepsgericht én organisatiegericht! / *Drs. L. Tijmens en dr. R. Poell*

Het vak van trainer (3/2002)

Gastredactie: A. Cnossen

- Het vak van trainer / *A. Cnossen en drs. S. Tjepkema*
- Trainers over hun vak (1): Outdoortraining / *E. Kalisvaart*
- De kunst en kunde van het zelfsturen ... en hoe je als trainer een bijdrage kunt leveren aan de ontwikkeling hiervan / *I. Khalil*
- Trainers over hun vak (2): E-learning / *W. Rubens*
- Persoonlijke ontwikkeling: een opleidingsnoodzaak. Achtergronden en aandachtspunten voor trainers / *H. Nathans*
- Trainers over hun vak (3): Video Interactie Analyse / *C. Smit*
- Training voor transformatie als uitdaging: een voorbeeld uit de praktijk van gendertrainingen / *H. Walters*
- Trainers over hun vak (4): RET / *T. Ijzermans*
- Spelen met de werkelijkheid: over het werken met acteurs in trainingen / *J. van der Steen*
- Gaming: een krachtig leermiddel / *Prof. dr. L. de Caluwé en drs. A. Stoppelenburg*
- Trainers over hun vak (5): Imaginatie / *J. Taal*
- Bouwen aan het vak en aan vakbekwaamheid. Een schets van de professionele ontwikkeling van trainers en hun vakgebied met implicaties voor train-de-trainer trajecten / *K. van Lohuizen*
- Structurele professionalisering van trainers: het studiehuis van SBI / *M. Sybesma*
- Onderwijsvernieuwing en gedragsverandering bij docenten. Het Droste-effect als voorwaarde voor succesvolle implementatie / *H. Sysling*
- Trainers over hun vak (6): Zingeving / *G. Kappers*
- Trainer als tweede beroep / *K. van Unen*
- Ignorance or arrogance, of waarom gaan ontwikkelingen voorbij aan het vak 'training'? / *M. Biesheuvel*

SUGGESTIES VOOR *HRD Thema*

Als u suggesties hebt voor nieuwe onderwerpen of voor verbetering van de nummers van *HRD Thema*, laat ze ons dan weten via dit formulier. We stellen uw mening en ideeën zeer op prijs.

Naam organisatie:

Naam:

Functie:

Adres:

Postcode/Plaats:

Telefoon:

Suggesties/opmerkingen:

U kunt dit formulier in een envelop zonder postzegel sturen naar:

Kluwer bv

Business Unit Human Resources Management

Antwoordnummer 10156

2400 VB Alphen aan den Rijn.

