

O De assessmentdriehoek voor leerwegaafhankelijk toetsen en begeleiden

Dit artikel is het drieënzeventigste in een serie praktische artikelen over onderwijsinnovatie. Deze serie heeft de bedoeling om mensen die werkzaam zijn in het hoger onderwijs handreikingen en aandachtspunten te bieden voor eigen initiatieven in onderwijsinnovatie. De onderwerpen van deze reeks kunnen uiteenlopen, maar zullen altijd gaan over het maken van onderwijs en dus over toepassingen van onderwijskundige en onderwijstechnologische inzichten in het dagelijks werk van de docent, het onderwijsteam of de opleidingsmanager.

Auteurs

Antoinette van Berkel

Van Berkel werkt bij de Hogeschool van Amsterdam en ondersteunt opleidingen op het gebied van toetsen en beoordelen, met als speciaal aandachtsgebied assessments en leerwegaafhankelijk toetsen. Daarnaast publiceert ze over genoemde onderwerpen. Meer informatie: <https://www.linkedin.com/in/antoinette-van-berkel-06604a10/>

Door gebrek aan ruimte is de lijst met bronnen bij dit artikel komen te vervallen. Deze is op aanvraag verkrijgbaar: a.van.berkel@hva.nl. Ook reacties op dit artikel kunnen naar dit adres.

Inhoud

- _ Samenvatting
- _ Inleiding
- _ De assessmentdriehoek
- _ Leerwegaafhankelijk toetsen en beoordelen met de assessmentdriehoek
- _ Professionaliseren en certificeren van assessoren
- _ Leren en begeleiden met de assessmentdriehoek
- _ De assessmentdriehoek als rode draad

Box 1: De assessmentdriehoek

Box 2: Voorbeeld van kerntaak uitgewerkt in beroepshandelingen

Box 3: Voorbeeld van uitgewerkt gedragscriterium op eindniveau

Box 4: Voorbeeld van rijke leeruitkomst in de assessmentdriehoek

Box 5: Voorbeelden van vragen over de elementen en verbindingsvragen

Box 6: Randvoorwaarden voor een valide en betrouwbaar leerwegaafhankelijk assessment

Samenvatting

Flexibel opleiden met leeruitkomsten is actueel in het hoger beroepsonderwijs (OCW, 2015). Leerwegaafhankelijke leeruitkomsten bieden studenten ruimte voor verschillende leeractiviteiten om beoogde resultaten te behalen. Zo ontstaat meer flexibiliteit in opleiden. Zogenaamde 'rijke' leeruitkomsten bevatten de ingrediënten van beroepsbekwaamheid en geven studenten bij het leren in de beroepspraktijk een kader om beroepsrelevante taken uit

te voeren en professioneel gedrag op hbo-niveau te ontwikkelen. Een generiek model om rijke leeruitkomsten te ontwerpen en leerwegaafhankelijk te toetsen is de assessmentdriehoek, met als herkenbare elementen: A) kerntaken representatief voor het beroep, B) professioneel gedrag voor hbo'ers, en C) bewijs, waarmee de student dit gedrag kan aantonen. Door de elementen van de driehoek structureel met elkaar te verbinden ontstaat een compleet beeld van de prestatie van de student (de leeruitkomst). Om recht te doen aan de student zijn deskundige begeleiders en assessoren nodig die de kunst van het werken met de driehoek verstaan.

Inleiding

Het hbo leidt studenten op tot beroepsbekwame professionals die voldoende toegerust zijn om te handelen in situaties die representatief zijn voor de kerntaken van het beroep waarvoor wordt opgeleid. De beroepspraktijk levert echter voortdurend nieuwe praktijkvraagstukken op die maken dat we het vaker 'niet weten'. Leren in het hbo vraagt erom dat studenten, docenten en de praktijk samen optrekken. Het is een zoektocht waarbij docenten handreikingen doen aan studenten om praktijkvragen scherp te stellen, strategieën en modellen toe te passen en tot originele analyses en oplossingen te komen. Het is niet voor niets dat onderzoekend vermogen de laatste jaren aan terrein heeft gewonnen. De

vraag is hoe we ervoor zorgen dat studenten de juiste dingen leren om beroepsbekwaam te worden en zich staande te houden in een continu veranderende beroepspraktijk. Een hulpmiddel is de assessmentdriehoek (Van Berkel & Te Lintelo, 2007). Deze laat ruimte voor diversiteit in praktijktaken en -situaties en biedt tegelijkertijd een kader voor het ontwikkelen van professioneel gedrag. Het is een beproefd instrument om rijke leeruitkomsten - gericht op professioneel handelen bij het uitvoeren van diverse beroepstaken in verscheidene praktijksituaties - te ontwerpen en studenten te begeleiden en te beoordelen. Dit artikel begint met een toelichting op de afzonderlijke elementen van de assessmentdriehoek, met oog voor kwaliteitscriteria die gelden voor de uitwerking van elk van de elementen. Ook het duiden van niveau komt aan bod. Daarna volgt uitleg over leerwegaafhankelijk toetsen en beoordelen met behulp van de assessmentdriehoek en de rol van de assessor daarin. Vervolgens is er aandacht voor de toetsbekwaamheid van assessoren en de wijze waarop zij daaraan kunnen werken in een professionaliserings- en certificeringsprogramma. Tot slot komt de leerfunctie van de assessmentdriehoek aan de orde.

De assessmentdriehoek

De assessmentdriehoek (box 1) geeft

houvast bij het ontwerpen van rijke leeruitkomsten die gaan over professioneel gedrag bij het uitvoeren van kerntaken in uiteenlopende beroepssituaties en -contexten. De drie elementen van de driehoek en de verbinding daartussen presenteren een rijke leeruitkomst. Het betreft:

- A. Kern- of beroepstaken representatief voor de beroepspraktijk;
- B. Professioneel gedrag, afgeleid van de hbo-standaard;
- C. Bewijs in de vorm van beroepsproducten, gedrag en verantwoording.

A. Kern- of beroepstaken

Kern- of beroepstaken omvatten een scala aan professionele activiteiten, zoals praktijkopdrachten en -vraagstukken, beroepshandelingen, dilemma's, problemen en kritische situaties. Ze zijn relevant en representatief voor de toekomstige professie of de beroepspraktijk waarvoor wordt opgeleid en ze dragen bij aan het realiseren van de eindkwalificaties van de opleiding. Kerntaken zijn te vinden in het opleidingsprofiel en bij elkaar omvatten de kerntaken 'wat' de student moet doen om de kneepjes van het vak te leren. Afhankelijk van het moment in de opleiding dat een student met een specifieke taak aan de slag gaat, krijgt de uitwerking ervan vorm in de driehoek. Daarbij geldt:

hoe vroeger in de opleiding en hoe minder (praktijk)ervaring, hoe minder complex de taak en hoe meer handvatten de student nodig heeft om deze uit te voeren. Dit impliceert per kerntaak een uitwerking in samenhangende beroepshandelingen of deeltaken en soms ook in dilemma's of kritische situaties om de student (en begeleidend docent) houvast te geven (zie box 2). Een deeltaak of beroepshandeling, zoals een indicatie of diagnose stellen in het voorbeeld (box 2), kan zo complex of elementair zijn, dat onderwijs hierop wordt aangeboden om studenten de benodigde kennisbasis en vaardigheden bij te brengen. Dit gebeurt meestal in de eerste jaren van de opleiding. Naarmate de student vordert in de opleiding wordt meer zelfstandigheid van hem verwacht bij het uitvoeren van vergelijkbare taken. Element A van de driehoek wordt dan minder geëxpliciteerd en bestaat bijvoorbeeld uit een complexe en omvangrijke taak, een opdrachtbeschrijving of een schets van een praktijkvraagstuk.

Variatie in taken

Opleidingsprofielen bevatten een representatieve selectie van taken die relevant zijn voor de beroepspraktijk. Het feit dat het hier om een - weliswaar verantwoorde - verzameling gaat, biedt tegelijkertijd

ruimte om studenten keuzes te laten maken uit (vergelijkbare) taken passend in het profiel en representatief voor de eindkwalificaties van de opleiding. Het is natuurlijk afhankelijk van de flexibiliteit van een opleidingsprogramma, hoeveel er voor studenten te kiezen is. En het is goed denkbaar dat bepaalde taken zo belangrijk zijn voor het beroep dat elke student daarmee ervaring moet opdoen. Maar in elke opleiding zijn er momenten dat studenten keuzes maken, bijvoorbeeld bij stages en bij het afstuderen. Dan dient de vraag zich aan of een stageopdracht of praktijkvraagstuk wel in het profiel past? Gebleken is dat een overzicht van professionele taken die geschikt zijn voor een stage heel handzaam is voor alle betrokkenen. Het gaat dan om taken die in de praktijk goed kunnen worden uitgevoerd, die aansluiten bij het niveau van studenten in de opleiding én voldoende uitdagend zijn. Een lijst met taken en bijbehorende spelregels biedt studenten een kader bij het zoeken van een stageopdracht en -praktijk, passend bij hun individuele interesse, talent en profilering. Vanuit het perspectief van de opleiding is de lijst een hulpmiddel om grip te houden op relevante stage-opdrachten van niveau.

Het formuleren van representatieve kern- of beroepstaken vraagt om vakinhoudelijke en theoretische expertise, praktijkkennis en -ervaring. Het werkt goed om drie à vier docenten, eventueel onder begeleiding van een onderwijs- of toetskundige, voor hun vakgebied de vertaalslag te laten maken van complexe, abstracte en gedetailleerde beschrijvingen in het opleidingsprofiel naar herkenbare taken die voldoende ruimte laten voor variatie in praktijksituaties en werkwijze. Vragen die daarbij leidend zijn: Welke taken kunnen studenten op niveau X uitvoeren? Wat houden deze taken precies in? Kan de student deze taken zelfstandig uitvoeren? Een werkbaar lijst van taken bestaat uit een overzichtelijk aantal samenhangende (deel)taken, die voldoende onderscheidend zijn van elkaar. Kerntaken zijn

BOX 2: VOORBEELD VAN KERNTAAK UITGEWERKT IN BEROEPSHANDELINGEN

Kerntaak: Het behandelen van cliënten bij complexe hulpvragen (het betreft een taak op eindniveau in een paramedische opleiding).

Beroepshandelingen zijn:

1. Onderzoeken van hulpvraag, bijvoorbeeld via meetinstrumenten, gesprekstechnieken of richtlijnen.
2. Screenen, indicatie en diagnose stellen (inclusief consult).
3. Ontwerpen van behandelplan of programma voor individu en/of groep.
4. Uitvoeren van behandeling of programma met behulp van erkende methodes of vanuit (para)medische richtlijnen.
5. Evalueren van een behandeling of programma: kwaliteit vaststellen op basis van meetbare effectiviteit, efficiëntie en cliëntgerichtheid.
6. Verslagleggen van behandeling volgens professionele (para)medische richtlijnen.
7. Beslissingen nemen op grond van de beschikbare evidence: toepassen van resultaten uit extern bewijs.
8. Evalueren van de kwaliteit van het proces: overdragen aan collega's en opslaan voor toekomstig gebruik.
9. Delen van kennis en expertise met vakgenoten, via het geven van werkbegeleiding, het deelnemen aan intervisiebijeenkomsten of intercollegiaal overleg.
10. Ontwikkelen van professionaliteit via klinisch redeneren en reflecteren.

Body of Knowledge and Skills: Voor het uitvoeren van deze kerntaak is actuele kennis nodig van concepten en modellen op het gebied van Evidence Based Practices, werken in teams, communicatiestrategieën, projectmanagement en zorgfinanciering.

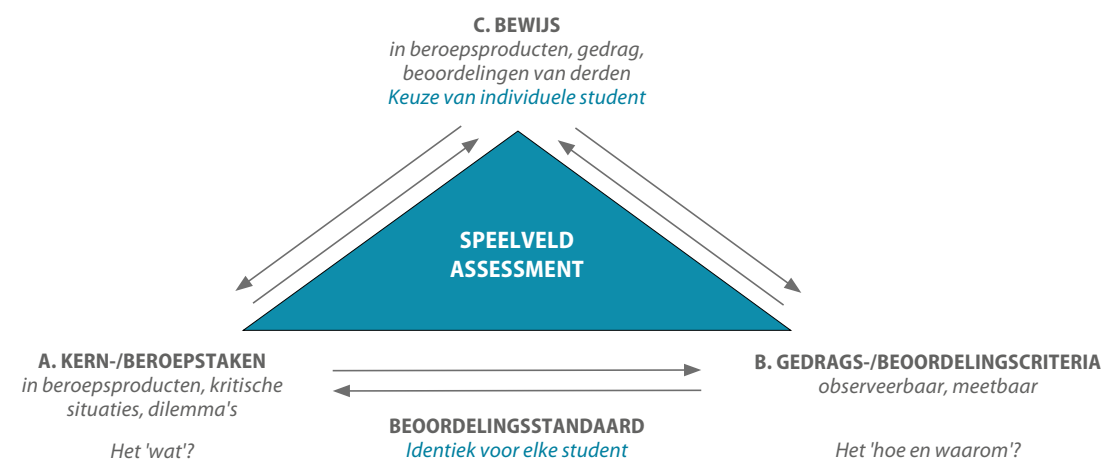
ook duurzaam en hoeven bij een wijziging in het curriculum niet direct vervangen te worden. Als proef op de som is het aan te raden om in een vroeg stadium vertegenwoordigers uit de beroepspraktijk en studenten te betrekken bij het uitwerken van de taken en hen om feedback te vragen. Zij gaan er tenslotte mee aan de slag in de praktijk. Samengevat voldoen kerntaken aan de volgende kwaliteitscriteria (gebaseerd op NVAO, 2015). Ze zijn *representatief* voor de eindkwalificaties van de opleiding, *herkenbaar* voor studenten, (praktijk)begeleiders en assessoren, *generiek* (bieden ruimte voor variatie in beroepspraktijken, -situaties en werkwijze), *transparant* (overzichtelijk, hanteerbaar, samenhangend en onderscheidend van elkaar), *duurzaam* (gaan jaren mee), en *realistisch* (afgestemd op het niveau van de student in de opleiding).

B. Professioneel gedrag

Rechtsonder in de driehoek staat het gedrag dat van beroepsbekwame hbo'ers wordt verwacht bij het uitvoeren van de

kerntaken. Het gaat dan over vragen als: Hoe pakken studenten de kerntaak aan? (onderzoekend, door kennis toe te passen en maatschappelijk bewust) en waarom op die manier? (verantwoording). De onderzoekende en kritische houding die hieraan ten grondslag ligt gaat gepaard met het professionele gedrag dat studenten als 'tweede natuur' dienen te ontwikkelen en waarop zij worden beoordeeld. De gedragscriteria zijn daarom de beoordelingscriteria bij het toetsen en beoordelen van leeruitkomsten. Ze geven antwoord op de vraag: Wanneer vinden we het uitvoeren van een kerntaak goed genoeg? De gedragscriteria zijn generiek en afgeleid van de hbo-bachelorstandaard die staat voor het hbo-bachelorniveau: 1. (toepassen van) een stevige kennisbasis, 2. onderzoekend vermogen, 3. professioneel vakmanschap, en 4. maatschappelijk en ethisch verantwoord handelen (HBO-Raad, 2009). Het valt direct op dat deze duidingen te abstract zijn voor studenten om het gewenste gedrag te ontwikkelen, voor docenten om die ontwikkeling te begeleiden en voor assessoren

BOX 1: DE ASSESSMENTDRIEHOEK



om het gedrag te toetsen en beoordelen. Hier moet een slag overheen om het concreet, meer specifiek (voor opleiding en beroep) en meetbaar te maken. Voorbeelden van concreet gedrag zijn te vinden in het opleidingsprofiel, hoewel niet altijd zo gelabeld of gerubriceerd. Bij voorkeur wordt het beroepsspecifieke gedrag uit het profiel herkenbaar ondergebracht in de vier gedragscategorieën van de hbo-standaard, zodat deze betekenis krijgen voor studenten, begeleiders, assessoren en werkveld. Onderdeel hiervan is het duiden van concepten, modellen, protocollen, etc. in een zogenaamde Body of Knowledge and Skills (BoKS), die voor elke kerntaak specifiek is. Een voordeel van de voorgestelde werkwijze is dat bij accreditatie in één oogopslag duidelijk is hoe het gedrag van de leeruitkomst zich verhoudt tot het opleidingsprofiel en tot de vier genoemde onderdelen van de hbo-standaard. De kunst is om het gewenste gedrag te vervatten in een hanteerbare set van criteria die de lading dekt. Hanteerbaar betekent hier 'behapbaar': niet te veel en niet te weinig, niet te gedetailleerd en niet te open.

Gedragscriteria zijn een hulpmiddel om het oordeel van de deskundige te onderbouwen. In dit verband is het goed te beseffen dat professioneel gedrag niet te vangen is in gedetailleerde (afvink)lijsten (Sluijsmans, 2013), noch in rubrics, al zijn deze laatste wel een handig hulpmiddel om gedrag te observeren, te bevragen en te toetsen. Met een team docenten, eventueel onder begeleiding van een onderwijskundige, gezamenlijk rubrics ontwikkelen, is een manier om tot de kern van het vereiste gedrag te komen én tot een gemeenschappelijke interpretatie daarvan. Het draagt bij aan de validiteit en betrouwbaarheid van de rubric en maakt het gebruik ervan zinvol. Aanvullend kan een overzicht met voorbeeldgedrag per gedragsindicator als onderlegger dienen, om de interpretatie te borgen. En als hulpmiddel voor studenten om het eigen gedrag te herkennen of om feedback te geven aan medestudenten (zie box 3). De genoemde kwaliteitscriteria voor de kerntaken zijn ook van toepassing op de gedragscriteria. Enkele daarvan krijgen een iets andere invulling en het criterium 'toetsbaar' is aangevuld. Gedragscriteria

(gebaseerd op NVAO, 2015) zijn: *representatief* (afgeleid van de hbo-standaard en het opleidingsprofiel), *open en generiek* (toepasbaar bij meerdere kerntaken en in verschillende beroepssituaties), *toetsbaar* (uitgewerkt in een hanteerbare set van specifiek, meetbaar en observeerbaar gedrag), en *transparant* (de relatie met de (eind)kwalificaties van de opleiding is duidelijk). Bij het formuleren van zowel de kerntaken als de gedragscriteria speelt taalgebruik, waaronder jargon, een belangrijke rol. De opleiding en de beroepspraktijk moeten elkaar immers goed verstaan. Een reden te meer om docenten te betrekken bij de uitwerking van de criteria en zowel vertegenwoordigers van het werkveld als studenten van de opleiding kritisch mee te laten lezen (Sluijsmans, 2013). Bij het toetsen en beoordelen van leeruitkomsten vormen de kerntaken A en de gedragscriteria B de beoordelingsstandaard (A <-> B, zie box 1). Deze is identiek voor elke student met ruimte voor verschillende (praktijk)opdrachten en -situaties, aanpak en leeractiviteiten. Dat maakt dat de standaard ook leerwegaanhanke-

lijk kan worden ingezet. Niveau aanbrenge- gen in de standaard is een lastige klus. Het heeft immers betrekking op zowel de kerntaken en de context waarin deze worden uitgevoerd als op de gedragscriteria. Het opleidingsprofiel geeft meestal aanknopingspunten om niveau te duiden, maar die zijn doorgaans abstract en multi-interpretabel. Bulthuis (2013) hanteert de dimensies 'zelfstandigheid' en 'complexiteit' van taak en context: hoe zelfstandiger de student kerntaken uitvoert en hoe complexer die taken zijn, hoe hoger het niveau. Bij de mate van zelfstandigheid valt te denken aan meer of minder begeleiding, sturing, interactie of vrijheid in het maken van keuzes bij de uitvoering en afbakening van een taak. Complexiteit is tweeledig en gaat zowel over taak als context. De taak wordt ingewikkelder als bijvoorbeeld een voor de hand liggende oplossing ontbreekt of als de geleerde standaardprocedure niet van toepassing is. De complexiteit van de context neemt onder andere toe in situaties waarin rekening moet worden gehouden met verschillende opdrachtgevers en betrokkenen en/of met tegenstrijdige belangen. De dimensies bieden handvatten en speelruimte om als team met elkaar in gesprek te gaan over het gewenste niveau en om hiervoor een norm te ontwikkelen. Ook de SOLO (Structure of Observed Learning Outcomes) taxonomie van Biggs & Collis (1982) leent zich voor het aanbrenge- van niveaus in gedrag (B). In deze taxonomie, bestaande uit vijf niveaus, gaat een hoger kennis- en begripniveau vooraf aan een hoger niveau van professioneel handelen. Een absolute duiding van niveau lijkt niet realistisch. Daarvoor is de praktijk te divers en de invloed daarop vanuit de opleiding is vrijwel nihil. Wel kan de opleiding minimale eisen stellen aan de complexiteit van kerntaken en de mate van zelfstandigheid bij de uitvoering ervan, wanneer studenten zelf opdrachten verwerven in de beroepspraktijk. Verder spreekt het voor zich dat studenten, begeleiders en assessoren met elkaar in gesprek moeten blijven over een gemeenschappelijk begrip van de be-

oordelingsstandaard en het vereiste niveau. Ook is het aan te raden om met enige regelmaat voorbeelden van praktijkopdrachten, ervaringen met de standaard en beoordelingen kritisch met elkaar door te nemen om het niveau beter in de vingers te krijgen.

C. Bewijs: beroepsproducten, gedrag en verantwoording

Het derde punt van de driehoek heeft betrekking op bewijs dat de student aanlevert. Het is het resultaat van de kerntaak die hij heeft uitgevoerd in de praktijk en het heeft de vorm van een beroepsproduct. Dit product dient als middel om met de student in gesprek te gaan over zijn professionele handelen bij de uitvoering van een kerntaak. Bewijs is bij voorkeur voorzien van een toelichting op de situatie, de eigen rol, werkwijze, resultaat en verantwoording daarvan en kan zowel een tastbaar product zijn als gedrag, zoals het toedienen van een injectie in paramedische opleidingen. Losse (2012) onderscheidt vijf types beroepsproducten, die het resultaat zijn van een systematische aanpak waarvan onderzoeksvaardigheden deel uitmaken: advies, ontwerp, artefact, handeling en onderzoek. Het zal duidelijk zijn dat het resultaat van het werken aan een kerntaak in de opleiding, bijvoorbeeld aan de hand van een opdracht of casus, voor vrijwel elke student hetzelfde resultaat zal opleveren. Echter, een vergelijkbare opdracht die studenten voor verschillende opdrachtgevers in de praktijk doen, levert een diversiteit aan (deel) producten op. Het resultaat is immers afgestemd op de wensen van de opdrachtgever. Dit stelt studenten in staat om zich te profileren aan de hand van uitgevoerde opdrachten in uiteenlopende contexten en de beroepsproducten die daarvan het resultaat zijn. De elementen van de driehoek staan niet op zichzelf. Geïntegreerd vormen ze een rijke leeruitkomst, waarbij steeds één kerntaak, met gerelateerde deeltaken, centraal staat (box 4). De driehoek in zijn geheel is het speelveld bij het uitvoeren

van die kerntaak in de praktijk (studenten), bij het begeleiden ervan (docenten, praktijkbegeleiders) en bij het toetsen en beoordelen (assessoren). De standaard - de kerntaak en de set van gedragscriteria in de onderkant van de driehoek - is richtinggevend voor de leeractiviteiten van de student en vormt als het ware een anker. Het spreekt voor zich dat de criteria zijn afgestemd op de kerntaak. Niet elke taak leent zich voor het ontwikkelen van al het gewenste gedrag. Zo kan het werken in een team slechts worden ontwikkeld als de student in teamverband aan een taak werkt. Maar er is een harde kern van gedragscriteria die altijd van toepassing is, zoals het werken vanuit een brede kennisbasis met een kritische, onderzoekende houding en met oog voor maatschappelijk verantwoord handelen.

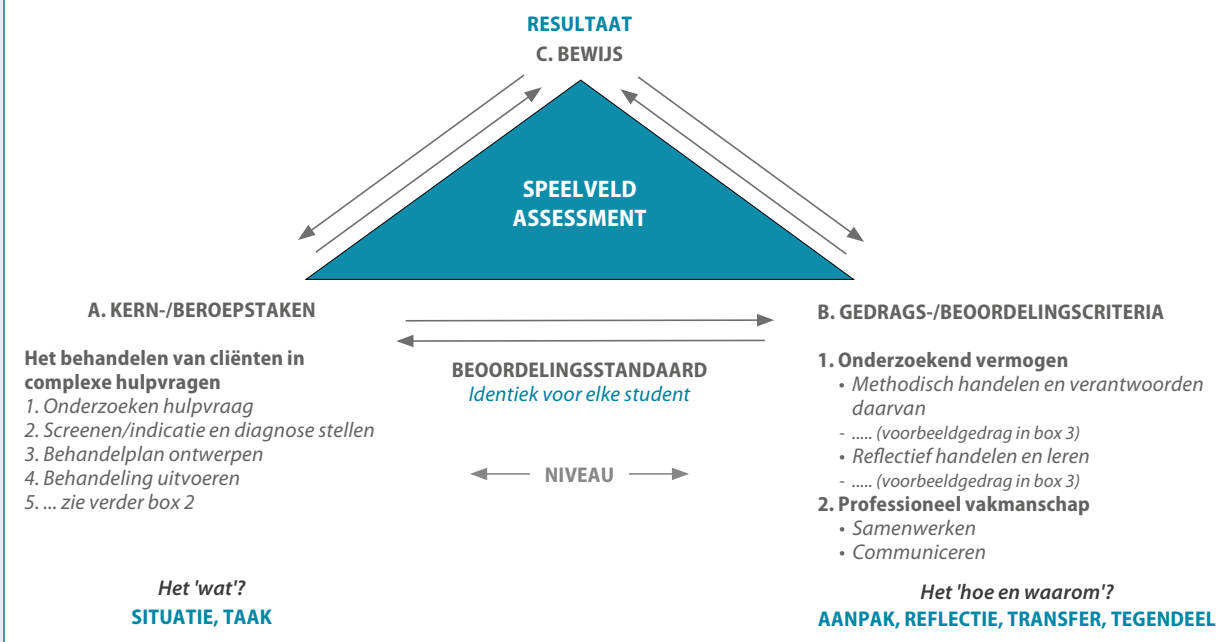
Leerwegaanhanke-lijk toetsen en beoordelen met de assessmentdriehoek

De rijke leeruitkomsten bieden studenten de ruimte om resultaten te boeken via verschillende leerwegen, -activiteiten en -situaties, bijvoorbeeld in de beroepspraktijk. Dat brengt veel variatie met zich mee in kern- of beroepstaken, kerncontexten, werkwijze en resultaten. Het zijn deze variabelen die een rol spelen bij het leerwegaanhanke-lijk toetsen van beroepsbekwaamheid. De open en generieke leeruitkomsten in de assessmentdriehoek bieden het kader: het beoordelingsmodel voor een valide en betrouwbare toets. Het doel van de toets is om een integraal beeld te krijgen van de prestaties (A <-> B) van de student aan de hand van bewijs (C). Volgens het principe van alignment (Biggs, 1999) zijn assessments een passende toetsvorm om informatie over de prestaties van studenten te verzamelen en interpreteren (Brookhart, 1999). Portfolio- en gedragsassessments, of een combinatie hiervan, komen daarom vaak voor in het hbo. Ook simulaties lenen zich voor het toetsen van kritische taken, maar hebben het nadeel dat de omgevingsfactoren weinig realistisch zijn, wat afbreuk doet aan de validiteit ervan.

BOX 3: VOORBEELD VAN 'ONDERZOEKEND VERMOGEN' UITGEWERKT IN GEDRAGSCRITE-RIA MET VOORBEELDDGEDRAG Eindniveau, paramedische opleiding

Hbo-standaard	Gedrags-/beoordelingscriteria	Voorbeeldgedrag (niet bedoeld om af te vinken!) De student:
Onderzoekend vermogen	Methodisch handelen en verantwoorden daarvan	<ul style="list-style-type: none"> - werkt systematisch: screenen, diagnosticeren, behandelplan, behandelen, evalueren en afronden; - werkt methodisch: hanteert actuele, relevante en bewezen modellen, concepten, (gespreks)technieken, protocollen en praktijken ('evidence-based practice', klinisch redeneren); - evalueert nieuwe richtlijnen, methoden of technieken; - verantwoordt diagnose, behandeling en werkwijze in overeenstemming met de vigerende beroeps- en gezondheidsstandaarden, de beroepscode en wetgeving; - toont inzicht in de consequenties van het eigen handelen voor de cliënt en andere betrokkenen.
	Reflectief handelen en leren	<ul style="list-style-type: none"> - ontwikkelt zich door zelfreflectie en zelfbeoordeling van eigen (leer)resultaten; - stelt zich open voor feedback en handelt hiernaar/neemt ontvangen feedback aan; - onderzoekt leemten in eigen competentiebeheersing en werkt op basis daarvan aan zijn professionele ontwikkeling; - toont verantwoordelijkheid voor zijn professioneel handelen; - toont zich bewust van de beroepscode en de beroepsrol en handelt hiernaar; - is zich bewust van discrepanties tussen eigen waarden en normen en professionele waarden en normen, en herkent effecten hiervan op het eigen gedrag en dat van de ander.

BOX 4: RIJKE LEERUITKOMST IN DE ASSESSMENTDRIEHOEK



Zowel portfolio- als gedragsassessments bestaan uit twee delen: het observeren van gedrag en het voeren van een assessmentgesprek, ook wel criteriumgericht interview genoemd. In een portfolio-assessment bestaat de gedragsobservatie uit het bestuderen van het portfolio van de student met daarin bewijs van geleverde prestaties en een geschreven toelichting daarop. In een gedragsassessment observeren assessoren concreet gedrag van de student bij het uitvoeren van kritische taken in (een) authentieke beroepsituatie(s), bijvoorbeeld tijdens een stage. Daarna volgt in beide gevallen een assessmentgesprek, waarin de student zijn gedrag toelicht en gemaakte keuzes verantwoordt aan twee onafhankelijke en deskundige assessoren. Zij gaan als volgt te werk: per kerntaak stellen zij gerichte vragen over de elementen en zogenaamde verbindingsvragen om de elementen met elkaar te verbinden (zie pijlen in box 1 en 4). Ze kunnen vanuit elk element insteken, afhankelijk van de informatie die ze nodig hebben om een volledig en geïntegreerd beeld van de studentprestaties

te krijgen (zie box 5). Als de vragen en/of antwoorden geen betrekking hebben op de elementen van de driehoek en de verbinding daartussen - het speelveld van het assessment - leveren ze geen relevante informatie op om een oordeel te kunnen vormen over de gedragscriteria. De driehoek geeft naast een kader ook structuur aan het gesprek. Elke kerntaak wordt op een vergelijkbare manier uitgevraagd tot op het niveau van concreet gedrag dat de student heeft laten zien. Vanzelfsprekend zijn aanvullende technieken nodig om diepgang in het gesprek te brengen. Daarvan is de Situatie – Taak – Aanpak – Resultaat – Reflectie – Transfer – Tegendeel (STARRTT)-techniek een wezenlijke, omdat deze een directe relatie heeft met de elementen van de driehoek (zie box 4). Bovendien kunnen assessoren met deze techniek de reikwijdte van de beroepsbekwaamheid van de student verkennen, door te vragen naar Transfer (andere situaties waarin de student een vergelijkbare opdracht heeft uitgevoerd en zijn aangepaste aanpak daarbij) en Tegendeel (situ-

aties waarin het niet lukte om professioneel te handelen bij vergelijkbare opdrachten). Maar ook technieken om regie te houden (luisteren-samenvatten-doorvragen), te versnellen en het gesprek in goede banen te leiden (meta-communiceren), dragen bij aan het voeren van een goed assessmentgesprek. Zowel de vragen van het assessmentgesprek, als de verkregen antwoorden daarop, registreren de assessoren op een daarvoor bestemd protocol. Na afloop van het assessmentgesprek vormen zij eerst onafhankelijk van elkaar een oordeel op basis van de verkregen informatie uit het portfolio of de gedragsobservatie én het assessmentgesprek. Ze doen dat door per kerntaak na te lopen in welke mate de student heeft voldaan aan de vereiste gedragscriteria en kennen per kerntaak een score toe, bijvoorbeeld O, V, G, ZG, met behulp van beslisregels (Expertgroep protocol, 2014). Zoals: onvoldoende: de prestatie voldoet niet aan de gedragscriteria; voldoende: de prestatie voldoet niet op alle punten aan de gedragscriteria, maar in voldoende mate; goed: de prestatie vol-

BOX 5: VOORBEELDEN VAN VRAGEN OVER DE ELEMENTEN EN VERBINDINGSVRAGEN

Starten met één van de elementen van de assessmentdriehoek	Voorbeelden van vragen over de elementen
Kern- of beroepstaken	In je portfolio zien we dat je een <i>project hebt geleid</i> . Beschrijf eens in welke concrete situatie dat gebeurde? Wat was precies het probleem? En wat was jouw projectopdracht? Voor wie voerde je die uit? Wat was de omvang van het project? Wie waren erbij betrokken? Wat was jouw rol?
Beoordelingscriteria	Hoe ben je te werk gegaan bij het <i>scherp stellen</i> van de projectvraag? Wie heb je geraadpleegd? Welke literatuur heb je gebruikt? Aan wie heb je de – door jou geformuleerde projectopdracht – voorgelegd? Wat was zijn/haar feedback? Wat heb je met die feedback gedaan?
Bewijs	Wat vind je van <i>het projectplan in jouw portfolio</i> ? Hoe is het <i>gewaardeerd door je opdrachtgever</i> ? Hoe is het tot stand gekomen? Wie waren erbij betrokken? Welke literatuur/ bronnen/formats heb je gebruikt? Hoe ben je tot een keuze gekomen voor deze vorm en inhoud?
Verbinding leggen tussen de elementen van de assessmentdriehoek	Voorbeelden van verbindingsvragen
Verbinding A-B: van taak naar gedragscriteria	Hoe heb je <i>het project aangepakt</i> ? Welke stappen heb je <i>achtereenvolgens gezet</i> ? Welke methodiek gehanteerd? <i>Wat was er beschikbaar in de organisatie waarvan je gebruik kon/moest maken</i> ? Hoe ben je daarmee omgegaan? Als je <i>terug kijkt</i> op de aanpak die je voor dit project hebt gehanteerd, waarop ben je dan trots?
Verbinding A-C: van taak naar bewijs	Het resultaat van het project dat je in situatie X hebt geleid is het <i>projectplan in je portfolio</i> . Licht eens toe waarom je voor dit bewijs hebt gekozen? Wat precies was de <i>opdrachtbeschrijving</i> ? Hoe heb je het <i>mijlpalenplan</i> gehanteerd?
Verbinding B-A: van gedragscriteria naar taak	Leg ons eens uit waaruit blijkt dat je <i>onderzoekend bent geweest bij het managen van dit project</i> ? In welke fasen van het project was dat het geval? En waaruit bestond jouw onderzoekende houding?
Verbinding B-C: van gedragscriteria naar bewijs	We zijn op zoek naar de wijze waarop je de <i>hoofd- en deelvragen in het project hebt scherp gekregen</i> . Leg dat eens uit met behulp van het <i>projectplan</i> in jouw portfolio. Welke bronnen (<i>medewerkers</i>) en/of literatuur heb je geraadpleegd? Hoe vertaalt zich dat naar dit <i>projectplan</i> ?
Verbinding C-A: van bewijs naar taak	Kun je aan de hand van <i>dit projectplan</i> toelichten dat je daadwerkelijk een <i>project hebt geleid</i> en de <i>deeltaken die daarmee samenhangen hebt uitgevoerd</i> ?
Verbinding C-B: van bewijs naar gedragscriteria	Kun je ons in je <i>projectplan</i> laten zien voor welke onderdelen je <i>bronnen of literatuur hebt geraadpleegd</i> ? Op welke onderdelen heb je <i>eerder feedback gehad</i> ? En wat heb je met die <i>feedback gedaan</i> ? Tot welke <i>aanpassingen</i> in het plan heeft dat geleid?
In dit voorbeeld is de kerntaak een project leiden, het beoordelingscriterium onderzoekend vermogen, het bewijs van de student een projectplan. Voor de herkenbaarheid van deze elementen, zijn de verwijzingen ernaar in de voorbeeldvragen cursief gedrukt.	

doet op alle punten aan de gedragscriteria; zeer goed: de prestatie overtreft de gedragscriteria. Het gaat hier om een kwalitatieve beoordeling. In dit proces is het zogenaamde 'timmermansoog' belangrijk. De onderliggende vraag is steeds: Zie ik deze student als hbo-professional functioneren in de beroepspraktijk? Elke score die de assessor toekent, vergt een holistische kijk vanuit de eigen vak- en praktijkdeskundigheid, onderbouwd met observaties en

bevindingen uit het assessment die betrekking hebben op de gedragscriteria in de standaard. Dat is niet alleen nodig voor de objectiviteit van het oordeel, maar ook voor de legitimatie ervan naar de student. Na deze individuele stap vergelijken de assessoren hun scores en onderbouwingen. Ze gaan in gesprek over verschillen en overeenkomsten, waarbij ze elkaar inzicht geven in de wijze waarop ze tot de toegekende scores zijn gekomen. Om tot consensus te komen is het

nodig om dicht bij de standaard te blijven. Voor het omzetten van de kwalitatieve oordelen per kerntaak naar één eindbeoordeling, uitgedrukt in een cijfer, passen de assessoren gezamenlijk beslissingsregels toe. Deze regels bevatten marges voor het bepalen van het cijfer, bijvoorbeeld 'alle kerntaken voldoende: cijfer tussen 5,5 en 7,0'. Dat laat ruimte voor de deskundigheid van de assessoren om een onderbouwd cijfer te geven, waarmee zij recht kunnen doen aan de prestaties van

de student, bijvoorbeeld in termen van 'een krappe voldoende' (5,5) of een 'ruime voldoende' (7,0). De beslisregels maken deel uit van het beoordelingsformulier waarop assessoren het onderbouwde oordeel vastleggen, evenals het protocol waarop de assessoren hun voorbereiding, vragen, observaties en bevindingen van het assessment hebben genoteerd. Zowel het protocol als de beslisregels dragen bij aan de transparantie en herleidbaarheid van het oordeel. Na afloop van het beoordelingsproces geven assessoren het oordeel aan de student en onderbouwen dit met hun observaties en bevindingen, zodat het voor de student begrijpelijk en herkenbaar is. In box 6 zijn belangrijke randvoorwaarden voor valide en betrouwbare, leerwegaafhankelijke assessments op een rij gezet.

Professionaliseren en certificeren van assessoren

Hierboven is in grote lijnen beschreven hoe het leerwegaafhankelijk toetsen en beoordelen van rijke leeruitkomsten in zijn werk gaat. Dat vraagt om onafhankelijke en deskundige assessoren. Onafhankelijk impliceert dat assessoren geen andere relatie hebben met de student dan die welke nodig is voor het uitvoeren van het assessment. In de praktijk is meestal één van de assessoren niet betrokken geweest bij de begeleiding van de student. Deskundigheid kenmerkt zich allereerst door vakinhoudelijke expertise, brede kennis van de beroepspraktijk en praktijk-

ervaring. Maar om de kneepjes van het vak te leren en toetsbekwaam te worden is aanvullende professionalisering nodig (Andriessen, et al., 2017). Deze is gericht op een gemeenschappelijke interpretatie van de beoordelingsstandaard (A<->B), een eenduidige werkwijze bij het beoordelen van studenten, met behulp van een protocol, en het toepassen van passende gespreks- en beoordelingstechnieken. Het uiteindelijke doel is om studenten volledig tot hun recht te laten komen in een assessment. De Hogeschool van Amsterdam (HvA) heeft een assessorentraining van een dag ontwikkeld om dit doel te realiseren. De training is bedoeld voor docententeams. Om de training zinvol te laten zijn, is het een voorwaarde dat een opleiding beschikt over deugdelijke assessmentmaterialen, zoals een heldere beoordelingsstandaard en een werkbare assessmentprocedure. In de training, die de vorm van een simulatie heeft, doorlopen de deelnemers alle stappen van het assessment met een inhoudsdeskundige acteur in de rol van student en een authentiek portfolio als oefenmateriaal. De acteur heeft zich, evenals de deelnemers, vooraf verdiept in het portfolio. Herkenbare onderdelen in de training zijn: het filteren van informatie uit een portfolio ter voorbereiding op het assessmentgesprek via een protocol (voorbereiding); het bespreken van ieders bevindingen in de training (overeenkomsten en verschillen); het voeren van een assessmentgesprek met behulp van de assess-

mentdriehoek en toepassing van diverse gesprekstechnieken; het vormen van een onderbouwd oordeel; het geven van feedback en het omgaan met bezwaar. Deelnemers vinden vooral het oefenen van de gesprekstechnieken in een assessmentgesprek met de acteur (simulatie) en het gezamenlijk bespreken en toelichten van de beoordelingen (kalibreren) waardevolle elementen in de training. De individuele feedback die deelnemers krijgen op hun gespreksvoering geeft eye-openers en richting aan ieders eigen leerproces. Om deelnemers beslagen ten ijs te laten komen op de trainingsdag is een digitaal leerprogramma ingericht met kennisclips en filmpjes van good practices aangevuld met opdrachten, referentiewerken en literatuur. Deze hebben betrekking op de vaardigheden en het gedrag dat van assessoren wordt verwacht en dat is vastgelegd in een assessorprofiel (Van Berkel, 2016, pp. 4-5). De onderwerpen zijn: een goed assessmentgesprek voeren, (werken met) de assessmentdriehoek, (werken met) de STARRTT-techniek, een onderbouwd oordeel geven, feedback geven en omgaan met bezwaar en lastige situaties. Het materiaal is generiek, dus bruikbaar voor alle opleidingen. Deelnemers bereiden zich voor door het beeldmateriaal te bestuderen, soms aan de hand van een kijkvraag. Ook maken ze enkele praktische opdrachten, zoals het vullen van de assessmentdriehoek met taken, gedragscriteria en voorbeelden van bewijs uit de beoordelingsstandaard van de eigen opleiding en het doornemen van het portfolio dat tijdens de training wordt gebruikt volgens een vaste systematiek en met een protocol. De voorbereiding neemt ongeveer tweeënhalf uur in beslag en is tijden plaatsafhankelijk. De training bereidt assessoren voor op hun taak, maar problematiseert ook. Deelnemers ervaren hun sterke kanten en leerpunten. Daarmee gaan zij aan de slag in de praktijk van het toetsen en beoordelen van studenten. Examencommissies vragen steeds vaker om gecertificeerde assessoren, vooral bij het leerwegaafhankelijk

toetsen van eindkwalificaties. Bij de HvA kunnen assessoren opgaan voor certificering als zij het gedrag uit het assessorprofiel voldoende hebben ontwikkeld. Dat gedrag is gericht op de deskundige, eenduidige en consistente uitvoering van assessments. Voor certificering beoordeelt een onafhankelijke assessmentdeskundige de assessor in een proeve van bekwaamheid. Dat wil zeggen dat hij het gedrag van de assessor observeert tijdens de uitvoering van een authentiek assessment. Als de assessor aan alle gedragsindicatoren uit het profiel voldoet, volgt certificering. Na het behalen van het certificaat zijn het onderhouden van assessorvaardigheden en het bevorderen van een gezamenlijke interpretatie van de standaard terugkerende onderwerpen van professionalisering.

Leren en begeleiden met de assessmentdriehoek

De assessmentdriehoek geeft niet alleen houvast bij het leerwegaafhankelijk toetsen en beoordelen van professioneel gedrag. Voor studenten die dit gedrag gaan ontwikkelen en docenten die dat leerproces begeleiden biedt de driehoek ook handvatten. Hoe dit in zijn werk kan gaan, illustreert dit voorbeeld van een paramedische opleiding. Studenten lopen in het vierde jaar een praktijkstage van een half jaar. Uitgewerkte kerntaken (zoals die in box 2) dienen ter inspiratie om met een praktijkhouder in gesprek te gaan over stagewerkzaamheden. Deze kunnen variëren van het ontwikkelen van een strategie om meer patiënten te werven tot het oplossen van een knelpunt in de praktijk. In een intakegesprek verkent de student met de praktijkhouder wat de mogelijkheden zijn, hoe de student een bijdrage kan leveren aan de praktijk en hoe hij aan zijn leerdoelen (box 2 en 3) kan werken. Op basis hiervan schrijft de student een persoonlijk ontwikkelplan. Tijdens de stage komen studenten om de twee weken naar de opleiding voor intercollegiaal overleg, met als doel inzicht te krij-

gen in het eigen leerproces, dat proces te stimuleren én van elkaar te leren op een manier die ook in de beroepspraktijk gangbaar is. Ze doen dit in leergroepen van maximaal twaalf studenten (peers) begeleid door een docent. In het begin van de stageperiode zijn de stageopdrachten, kerntaken, leerdoelen en bijbehorende beroepsproducten, onderwerp van gesprek. Per stage-opdracht of kerntaak bespreken studenten met elkaar de achterliggende vraag en een oplossingsrichting. Ook de aanpak met bijbehorend gedrag (box 3) komen aan de orde. Later in de stage presenteren studenten in dit overleg hun praktijkopdracht en -ervaringen aan de hand van beroepsproducten in ontwikkeling. Bijvoorbeeld een behandelplan of een patiëntverslag, hun gekozen aanpak en gedrag. Zowel de peers als de begeleidend docent stellen hierover verdiepende vragen met behulp van de assessmentdriehoek en bespreken het vereiste niveau, in de opleiding onder andere geduid met: zelfstandig therapeutisch, evidence based en patiëntgericht handelen. Studenten krijgen zo een goed beeld van vragen die zij ook tijdens het assessment aan het einde van de stage kunnen verwachten. In het tweede deel van het overleg gaat de groep uiteen in (homogene) tweetallen om peerfeedback op elkaars producten en aanpak te geven met behulp van de gedragscriteria (box 3). Ter inspiratie en/of uitdaging bekijken studenten daarnaast beroepsproducten van peers die qua niveau, inhoud en kwaliteit verschillen van hun eigen werk. De praktijkbegeleider draagt zijn steentje bij door tijdens de stage feedback te geven op de werkwijze en het gedrag van de student. Hiervoor zijn speciale formulieren ontwikkeld. In het intercollegiaal overleg stelt de begeleidend docent kritische vragen en checkt hij onder andere of studenten het eens zijn met de (peer) feedback en of het niveau waarop zij in de praktijk 'werken' aan de maat is. Daarnaast geeft de docent tussentijds één keer feedback en feedforward op de portfolio's van de studenten met daarin beroepsproduc-

ten in wording en een toelichting met STARR. In dit intercollegiale overleg krijgen studenten een breed beeld van de beroepspraktijk. Zij kunnen bij elkaar de kunst afkijken om in de eigen stagepraktijk tot betere prestaties te komen. Ook de (peer)feedback en de feedforward (docent) is erop gericht de stage naar een hoger niveau te tillen. Bovendien blijkt het geven van effectieve feedback in grote mate inzicht te verschaffen in beoogde leerdoelen, waardoor studenten leren om beter en kritischer naar hun eigen werk te kijken (Van Bakel, 2016).

De assessmentdriehoek als rode draad

Veel opleidingen van de HvA werken met de assessmentdriehoek. In eerste instantie was het model bedoeld om beoordelingsstandaarden voor assessments te ontwikkelen en om assessoren op te leiden. Inmiddels kent de assessmentdriehoek meer toepassingen. De driehoek is een generiek model dat goed bruikbaar blijkt bij het ontwerpen van rijke leeruitkomsten, bedoeld voor studenten om professioneel gedrag te ontwikkelen bij het uitvoeren van kerntaken in de beroepspraktijk. Wanneer de elementen van de driehoek (kerntaken, professioneel gedrag en bewijs) zijn uitgewerkt volgens de geldende kwaliteitscriteria, vormen zij de rode draad voor het leren, begeleiden, leerwegaafhankelijk toetsen en beoordelen van studenten. Alle betrokkenen kunnen er hun voordeel mee doen: studenten, wanneer zij zich het gewenste professionele gedrag in de beroepspraktijk eigen maken; (praktijk)begeleiders, wanneer zij het leerproces van studenten begeleiden en hierop feedback en feedforward geven; assessoren, wanneer zij toetsen en beoordelen of studenten het vereiste gedrag laten zien bij het uitvoeren van kerntaken in uiteenlopende beroepssituaties. De praktijk heeft uitgewezen dat assessments aan kwaliteit, validiteit en betrouwbaarheid winnen, wanneer een goed uitgewerkte driehoek eraan ten grondslag ligt.

BOX 6: RANDVOORWAARDEN VOOR VALIDE EN BETROUWBARE LEERWEGAAFHANKELIJKE ASSESSMENTS

- Rijke leeruitkomsten: ontworpen volgens de systematiek en kwaliteitscriteria van de assessmentdriehoek.
- Gestructureerde en toegankelijke portfolio's (bij portfolio-assessments), met daarin bewijs van prestaties, zoals gevarieerde beroepsproducten met feedback en beoordelingen van verschillende beoordelaars (docenten, opdrachtgevers, klanten, praktijkbegeleiders, etc.).
- Een eenduidige en werkbare assessmentprocedure, inclusief protocollen, vastgelegd en toegankelijk voor alle betrokkenen.
- Onafhankelijke, theorie- en praktijkdeskundige assessoren.
- Toetsbekwame assessoren: opgeleid en bij voorkeur gecertificeerd.
- Twee assessoren per assessment/student voor een intersubjectief oordeel.
- Wisselende tweetallen van assessoren, ter voorkoming van 'nieuwe standaarden'.
- Kalibreersessies voor gezamenlijke interpretatie van beoordelingsstandaard, niveau en norm.