



onderwijs dat werkt

Toekomst Examinering Talen en Rekenen/wiskunde in het mbo

**Advies van Procesmanagement MBO 2010.
Januari 2008**

Colofon

Titel	Advies Examinering Talen en Rekenen/wiskunde in het mbo
Auteurs	Marianne Driessen, Karen Heij, Henk van der Kooij, Gerard Straetmans,
Versie	Versie 1.3
Datum	17 januari 2008

Inhoudsopgave

1 Aanleiding tot het advies	2
1.1 Waar komen de zorgen rondom de talen en rekenen/wiskunde vandaan?	2
1.2 Civiel effect van examinering	2
1.3 De opdracht	3
2 Examinering in het mbo: de specifieke situatie van talen en rekenen/wiskunde	5
2.1 De eindtermgerichte kwalificatiestructuur	5
2.2 De nieuwe competentiegerichte kwalificatiestructuur	6
2.2.1 Talen	6
2.2.2 Verplicht Nederlands in Leren, Loopbaan en Burgerschap	6
2.2.3 Eén verplichte mvt in LLB?	6
2.2.4 Rekenen/wiskunde	7
2.2.5 Verplicht rekenen/wiskunde in LLB?	7
2.3 Hoe is de borging van de kwaliteit van examinering nu vormgegeven?	7
3 Advies voor een examineringsmodel	10
3.1 Uitgangspunten	10
3.2 In gebruik zijnde varianten van examinering	12
3.2.1 Variant 1 Inburgeringsvariant	13
3.2.2 Variant 2 Voortgezet onderwijs variant	13
3.2.3 Variant 3 MBO-variant	13
3.2.4 Combinaties van varianten	13
3.2.5 Welke variant is het meest geschikt voor taal en rekenen/wiskunde in het mbo?	13
3.3 Advies voor examinering van taal- en rekentaalvaardigheid	16
3.3.1 Het examenmodel algemeen	16
3.3.2 Examenmodel voor Nederlands	18
3.3.3 Examenmodel voor de moderne vreemde talen (mvt)	20
3.3.4 Examenmodel voor Rekenen/wiskunde	22
3.4 Slagen of zakken?	24
4 Gestandaardiseerde examens in een modern jasje	26
4.1 Inleiding	26
4.2 Opzet en werking	26
4.2.1 Specificatie van te examineren kennis en vaardigheden	28
4.2.2 Opgavenproductie	28
4.2.3 Proefafname	28
4.2.4 Schaalconstructie en cesuurscores	29
4.2.5 Constructie examenversies	29
4.2.6 Vaststelling examenversies	30
4.2.7 Afname examens door instellingen	30
4.3 Computergestuurde adaptieve toetsing (CAT)	30

5	Aanbevelingen behorende bij dit advies	33
5.1	Bepaal de niveaus voor de talen opnieuw	33
5.2	Zorg voor goede taalassessoren	33
5.3	Taal mag geen struikelblok zijn	33
5.4	Stel minimale niveaus voor Rekenen/Wiskunde voor alle deelnemers in het kader van Burgerschap verplicht.....	34
5.5	Bepaal minimale niveaus voor Rekenen/Wiskunde per doorstroomvariant.....	34
5.6	Bepaal de vereiste niveaus voor Rekenen/Wiskunde per kwalificatieprofiel.....	34
5.7	Stel een beslismodel op.....	34
5.8	Betrek dit advies bij de ontwikkelingen van examenprofielen	34
	Literatuur	35
Bijlage 1	Onderzoeken naar taal- en rekenvaardigheden	37
Bijlage 2	Threshold Level.....	42
Bijlage 3	Raamwerk Rekenen/wiskunde	45
Bijlage 4	Inzet van schalingstechnieken voor examens	47

Inleiding

Voor u ligt een rapport met een advies over standaardisering van de examinering van de talen (Nederlands en de moderne vreemde talen) en rekenen/wiskunde in het middelbaar beroepsonderwijs. Dit rapport is tot stand gekomen in opdracht van het Procesmanagement MBO 2010 en dient als advies aan de Staatssecretaris van OC&W.

De groep van auteurs die het advies heeft opgesteld is samengesteld uit de volgende instituten met expertise rondom dit thema:

- Bureau ICE
- CINOP
- Cito
- Freudenthal Instituut

Het rapport is als volgt opgebouwd:

In hoofdstuk 1 wordt beschreven wat de aanleiding is geweest om te komen tot een advies over standaardisering van de examinering van taal- en rekenvaardigheid.

Hoofdstuk 2 gaat in op de problemen die er op dit moment spelen rondom examinering in het mbo in het algemeen en van de talen en rekenen/wiskunde in het bijzonder.

In hoofdstuk 3 wordt een advies gegeven voor een examineringsmodel met elementen van standaardisering.

Hoofdstuk 4 geeft een beeld van hoe gestandaardiseerde examinering 'in een modern jasje' er uit zou kunnen zien.

Tot slot worden er in hoofdstuk 5 enkele aanbevelingen gedaan die los staan van de ontwikkeling van examens zelf, maar er wel een duidelijke relatie mee hebben en soms zelfs voorwaardelijk zijn.

In de bijlagen vindt u uitgebreidere informatie over verschillende thema's die in de hoofdstukken aan de orde komen.

1 Aanleiding tot het advies

1.1 Waar komen de zorgen rondom de talen en rekenen/wiskunde vandaan?

Is het echt zo slecht gesteld met de beheersing van taal¹- en rekenvaardigheden van deelnemers in het mbo? Is het probleem nieuw en gekoppeld aan competentiegericht onderwijs of bestaat het al veel langer?

Het is snel en makkelijk gezegd dat onderwijsdeelnemers 'niet goed (meer) kunnen rekenen' of 'slecht(er) zijn (geworden) in taal'. Bij goed doorvragen naar wat er exact ontbreekt aan kennis en vaardigheid zijn antwoorden vaak weinig concreet. En als je doorvraagt naar wat dan precies nodig zou zijn, is menigeeen het antwoord ook schuldig. En daar zit nu ook precies de moeilijkheid. Om te kunnen vaststellen of het niveau van taal en rekenen/wiskunde voldoende is, moeten we ten eerste kunnen beschikken over een objectieve, gestandaardiseerde meetlat op basis waarvan we kunnen vaststellen hoe het met de beheersing van taal en rekenen/wiskunde is gesteld. En ten tweede moeten we met elkaar een duidelijk beeld hebben van wat exact nodig is om tevreden te kunnen zijn. Kortom: Welke norm hanteren we? Wanneer is er sprake van voldoende beheersing?

De afgelopen jaren zijn er diverse rapporten verschenen die ons bewust maken van zowel het belang dat we moeten hechten aan een voldoende beheersing van taal en rekenen/wiskunde als van de problemen die er in bepaalde contexten zijn voor wat betreft die beheersing. In Bijlage 1 zijn de uitkomsten van de diverse onderzoeken op dit terrein opgenomen. Met name de onderzoeken die wijzen op de problemen in de doorstroom van deelnemers van de ene onderwijssoort naar de andere, zoals bijvoorbeeld van MBO naar de PABO, hebben geleid tot een breed ervaren ongerustheid over de niveaus van taal en rekenen/wiskunde. Omdat een mbo-diploma een drievoudige kwalificering inhoudt (en dus recht geeft op die doorstroom) zijn de geconstateerde problemen een begrijpelijke bron van zorg. De onderzoeken laten ook zien dat de problemen en zorgen al beginnen bij het basisonderwijs en zich voorzetten in het voorgezet onderwijs en het mbo. Het is duidelijk dat de zeer recente invoering van het competentiegerichte leren in het mbo niet de oorzaak van de problemen kan zijn, maar eerder een kans biedt om te gaan werken aan een oplossing.

1.2 Civiel effect van examinering

Zorgen uitspreken over het al dan niet voldoende beheersen van taal en rekenen/wiskunde door onderwijsdeelnemers, is een eerste belangrijke stap. Willen we vervolgens ook kunnen waarborgen dat het niveau van taal en rekenen/wiskunde in het mbo voldoende is, dan is het nodig te beschrijven wat er precies aan taal en rekenen/wiskunde beheerst zou moeten worden. Dit kun je beschrijven in doelen voor een bepaalde onderwijsvorm, zoals dat in het beroepsonderwijs de afgelopen tien jaar achtereenvolgens gebeurde in eindtermendocumenten en in kwalificatiedossiers. Daarmee

¹ Daar waar gesproken wordt over 'taal' en 'taalvaardigheid' wordt bedoeld: taalvaardigheid in het Nederlands en in een of meer mvt's. De focus ligt echter in het algemeen op Nederlands. Daar waar de focus op de mvt's ligt wordt dit expliciet vermeld.

is geborgd 'wat er geleerd moet worden'. Om ook het niveau van beheersing te borgen is het nodig zorg te dragen voor controle op 'hoe er gemeten wordt': voor programma's van toetsing en afsluiting en van procedures die de kwaliteit daarvan kunnen borgen. Een vorm van controle op examinering is onontbeerlijk om resultaten van onderwijs te kunnen borgen. Daarnaast is examinering in de praktijk bijzonder sturend voor het onderwijs: dat wat wordt geëxamineerd wordt geleerd en de manier van examineren beïnvloedt ook de manier van leren en onderwijzen.

Als we willen dat taal en rekenen/wiskunde respectievelijk voldoende en passend zijn, dan moeten we erop toezien dat deze onderdelen op de juiste wijze geëxamineerd worden.

Het civiel effect van een examen treedt alleen op als betrokken partijen vertrouwen hebben in inhoud, vorm en normering van een examen. Dit vertrouwen in de examinering in het beroepsonderwijs staat momenteel regelmatig onder druk. De deelnemers aan het beroepsonderwijs hebben er de afgelopen jaren meermalen terecht op gewezen dat doel, opzet en inhoud van het onderwijs voor hen niet altijd even duidelijk waren. Het bedrijfsleven, erop gebrand werknemers in dienst te kunnen nemen die beschikken over voldoende relevante competenties om te kunnen starten op de arbeidsmarkt, maakt zich regelmatig zorgen over inhoud en niveau van het beroepsonderwijs. De politiek als belangenbehartiger van zowel studenten als de arbeidsmarkt is als derde partij erg alert op de borging van kwaliteit van het beroepsonderwijs. Het vermoede verschil in de waarde van een diploma van de ene school ten opzichte van een diploma van een andere school, terwijl het over dezelfde opleiding gaat, is een bron van zorg. Ook de concrete invulling van het onderwijs zodanig dat de einddoelen gehaald kunnen worden, baart de politiek zorgen².

Het is dan ook niet verbazingwekkend dat met name vanuit de politiek druk wordt uitgeoefend om ervoor te zorgen dat de examinering in het beroepsonderwijs voldoende waarborgen biedt voor kwaliteit. Het is belangrijk dat de belanghebbenden, betrokken bij het beroepsonderwijs, vertrouwen hebben in het gegeven onderwijs en, als bewijs daarvan, in het verstrekte diploma. De waarde van het diploma moet onbetwist zijn.

Bij twijfel over het bereikte niveau van taal en rekenen/wiskunde van mbo-deelnemers, is het dus van belang om zowel naar de inhoud van het onderwijsprogramma te kijken als naar het bereikte niveau, vast te stellen door examinering.

1.3 De opdracht

De opdracht tot dit advies luidde:

Uitgaande van het feit dat er een wens/eis is tot meer standaardisatie in de mbo-examens: formuleer een advies met betrekking tot de examinering van de talen (Nederlands en de mvt's) en rekenen/wiskunde waarbij een vorm van standaardisering wordt toegepast.

In dit advies concentreren we ons op de stappen die gezet zouden moeten worden om in te toekomst te kunnen vaststellen of mbo-deelnemers voldoende taal en

² Hierbij moet erop worden gewezen dat sommige zaken door de media nogal worden uitvergroot en dat de berichtgeving niet altijd zorgvuldig is.

rekenen/wiskunde beheersen. We concentreren ons op de twee pijlers die daarvoor nodig zijn:

- de invoering van een objectieve, gestandaardiseerde meetlat voor kennis en vaardigheid op het gebied van (vreemde) talen en rekenen/wiskunde
- het vaststellen van normen voor beheersing, gekoppeld aan die meetlat.

In dit rapport analyseren we eerst hoe beide pijlers momenteel in het beroepsonderwijs zijn uitgewerkt. Vervolgens stellen we een model voor examinering voor dat het mogelijk maakt de beheersing van taal en rekenen/wiskunde te meten, gekoppeld aan diezelfde meetlat, en daarmee de vastgestelde normen te bewaken. Dit alles binnen de context van het competentiegerichte middelbaar beroepsonderwijs.

2 Examinering in het mbo: de specifieke situatie van talen en rekenen/wiskunde

In dit hoofdstuk bekijken we of en in hoeverre de beide pijlers onder de borging van kwaliteit van taal en rekenen/wiskunde gerealiseerd zijn in de huidige uitwerking van het beroepsonderwijs. We bekijken of en in hoeverre er beschreven is wat de inhoud van het onderwijs zou moeten zijn en we bekijken of en in hoeverre er een objectieve, gestandaardiseerde meetlat beschikbaar is voor examinering en hoe de examinering momenteel verloopt.

Omdat het beroepsonderwijs momenteel werkt met twee verschillende generaties beschrijvingen van onderwijsdoelen, bekijken we in paragraaf 2.1 de stand van zaken voor de eindtermgerichte kwalificatiestructuur en in paragraaf 2.2 de stand van zaken voor de competentiegerichte kwalificatiestructuur. De eerste paragraaf is echter alleen bedoeld om terug te kijken. Het advies richt zich verder alleen op de nieuwe situatie van de competentiegerichte kwalificatiestructuur.

2.1 De eindtermgerichte kwalificatiestructuur

In de oude kwalificatiedossiers was er sprake van een lijst met eindtermen. Voor de moderne vreemde talen (mvt's) waren die nogal vaag omschreven, zoals bijvoorbeeld: 'Moet kunnen communiceren met klanten en leveranciers in mvt1 en mvt2'. Dit soort omschrijvingen zorgde voor twee problemen. Ten eerste was niet echt omschreven hoe goed iemand dat zou moeten kunnen. Ten tweede waren de eerste en de tweede mvt (en soms zelfs nog een derde) van dezelfde eindtermen voorzien waardoor het vrijwel onmogelijk werd om bijvoorbeeld Spaans net zo goed af te sluiten als Engels. Met betrekking tot de examinering moest er dus creatief worden omgegaan met die eindtermen. Een norm voor beheersing, gekoppeld aan een objectieve, gestandaardiseerde meetlat, ontbrak. En tenslotte waren er vaak géén specifieke eindtermen voor Nederlands, behalve daar waar Nederlands een beroepsgericht vak was, zoals bij secretariële opleidingen. Voor het vak Nederlands ontbrak op die manier zowel de inhoud (het 'wat') als de norm ('hoe goed'). Als gevolg hiervan kon er geen sprake zijn van een betrouwbare en/of valide beoordeling of examinering van de talen in het mbo.

Voor rekenen/wiskunde was de situatie vergelijkbaar. In de sector Techniek was sprake van een gedifferentieerde beschrijving van de eindtermen rekenen/wiskunde. Voor de beroepskwalificatie werden de vaardigheden en kennis alleen impliciet genoemd in het kader van de beroepstaken (en dat gold ook voor de andere sectoren), maar voor de doorstroomkwalificatie werd veelal een gedetailleerde lijst van specifieke eindtermen opgesomd. Dit was met name het gevolg van de inspanningen die in het TWIN³-project zijn gepleegd om een Techniek-breed programma te beschrijven dat door het HBO werd geaccepteerd als goede voorbereiding voor een studie binnen de sector Techniek (TWIN, 1998). In de jaren 2000-2004 zijn landelijke examens wiskunde en natuurkunde afgenomen op een beperkt aantal ROC's. Omdat het doorstroomprogramma niet werd

³ Techniek, Wiskunde, ICT, Natuurkunde was een curriculumproject (1996-2000) waarin een programma is gedefinieerd voor het basiscertificaat en voor het doorstroomcertificaat

opgenomen in het reguliere opleidingspakket, maar als extra belasting daar bovenop werd gezet en vanwege het feit dat een deelnemer zonder deze extra last ook toelaatbaar was in het HBO, nam de animo voor de centrale examinering snel af evenals voor het aanbieden van doorstroomgerichte vakonderwijs. Vanaf 2005 zijn er dus geen centrale examens meer afgenomen voor wiskunde. Er is binnen de eindtermengerichte kwalificatiestructuur geen enkel zicht op het examineren (al dan niet binnen de context van het beroep) van reken/wiskundige vaardigheden.

2.2 De nieuwe competentiegerichte kwalificatiestructuur

2.2.1 Talen

In de nieuwe competentiegerichte kwalificatiestructuur zien we een belangrijke verbetering ten opzichte van de eindtermengerichte kwalificatiedossiers voor wat betreft de te hanteren norm voor beheersing van taalvaardigheid. Alle dossiers bevatten beheersingsniveaus voor de talen, conform het door de Raad van Europa ontwikkelde *Common European Framework of Reference for Languages (CEF, of Europees Referentiekader voor de talen, ERK)*. Ieder kwalificatieprofiel bevat, naast de lijst met de beroepsgerichte competenties, ook nog een taalcompetentieprofiel voor Nederlands en voor de mvt's, voor zover nodig voor het beroep en daarmee dus gekoppeld aan de beroepscontext. De in de kwalificatiedossiers opgenomen taalcompetentieprofielen zijn bewust contextloos gelaten: er zijn alleen hokjes ingekleurd waarmee de normen voor niveaubehersing zijn aangegeven. De exacte inhoud, het 'wat', moet uit de contexten voor het beroep komen. Ook de beschrijving van de niveaus en de bijbehorende prestatie-indicatoren zijn bewust weggelaten. Deze zijn immers al beschreven en kunnen worden teruggevonden in het Europees Referentiekader voor de talen (ERK), het Raamwerk MVT (de Nederlandse bewerking voor het mbo) en het Raamwerk Nederlands (voor VMBO en MBO).

2.2.2 Verplicht Nederlands in Leren, Loopbaan en Burgerschap

Naast de taalcompetentieprofielen, gekoppeld aan het functioneren in een beroepscontext, zijn er in het document *Leren, Loopbaan en Burgerschap* ook verplichte eindniveaus Nederlands vastgesteld voor alle deelnemers, ongeacht de opleiding. Dit om te voorkomen dat deelnemers uitstromen met een te laag niveau Nederlands voor hun verdere leerloopbaan(doorstroom) en functioneren in de maatschappij (burgerschap). Het uitgangspunt hierbij is dat het hoogste niveau (ofwel het kwalificatiedossier ofwel het document *Leren, Loopbaan en Burgerschap*) geldt en bij de examinering moet worden aangetoond.

2.2.3 Eén verplichte mvt in LLB?

De Stuurgroep Competentiegericht Beroepsonderwijs heeft op 30 augustus 2007 aan de staatssecretaris geadviseerd over het opnemen van één verplichte mvt (gespecificeerd als Engels, Duits, Frans of Spaans) in alle mbo-opleidingen van niveau 3 en 4. Hierover (inclusief over het vereiste niveau voor die talen) is nog geen beslissing genomen. We gaan er van uit dat er een verplichting van één mvt wordt toegevoegd aan het document *Leren, Loopbaan en Burgerschap*.

2.2.4 Rekenen/wiskunde

Op dit moment is het zo dat rekenen/wiskunde vooral geïntegreerd is in de beroepsgerichte competenties en, net als bij de vorige generatie eindtermendocumenten, nergens expliciet beschreven is. Om die reden is er voor een expliciete beschrijving van rekenen/wiskunde zeer recent (2007) gewerkt aan een vergelijkbare aanpak als bij de beschrijving van taal door de ontwikkeling van het *Raamwerk Rekenen-wiskunde voor het mbo*. Dit beschrijvingskader lijkt goede mogelijkheden te bieden om voor een specifiek beroepsprofiel te omschrijven op welk vaardigheidsniveau (tussen volledig contextgebonden en abstract) één of meerdere van de reken/wiskundige domeinen dienen te worden beheerst. Deze vaststelling van vaardigheidsniveaus per kwalificatieprofiel moet echter nog gebeuren. In 2008 zal op verschillende niveaus (Kenniscentra en ROC's) de waarde van het instrument worden getoetst als beschrijvingskader van vereiste vaardigheden (in de Kwalificatiedossiers) en als instrument om onderwijs en toetsing vorm te geven. De huidige versie 0.9 van het *Raamwerk Rekenen-wiskunde* moet daarom ook worden gezien als een eerste pilot-versie.

2.2.5 Verplicht rekenen/wiskunde in LLB?

Ook voor rekenen/wiskunde wordt gedacht aan het verankeren van een minimaal niveau dat voor iedere mbo-gediplomeerde zou moeten gelden. In het *Raamwerk Rekenen-Wiskunde* wordt voor elk mbo-niveau een advies gegeven over de kennis en vaardigheden waarover een deelnemer minstens moet kunnen beschikken in het kader van Burgerschap. Nadrukkelijk moet worden gesteld dat dit niet de doorstroom betreft omdat die voor rekenen/wiskunde per kwalificatie kan verschillen. Op dit moment is er nog geen beslissing genomen over de minimale niveaus voor rekenen/wiskunde in het kader van Burgerschap. We gaan er echter in dit advies van uit dat de aanbevelingen uit het *Raamwerk Rekenen-Wiskunde* worden overgenomen en dat er verplichte minimale niveaus worden voorgeschreven.

2.3 Hoe is de borging van de kwaliteit van examinering nu vormgegeven?

De einddoelen van het beroepsonderwijs zijn per uitstroomrichting vastgelegd in eerder genoemde Kwalificatiedossiers. Dit gebeurt volgens een standaardformat. Zo is uniformiteit van het beschrijvingskader geborgd. Door de inhoud ervan op te stellen met betrokkenen uit onderwijs en arbeidsmarkt, is aansluiting van de inhoud op de vereisten van de arbeidsmarkt ook geborgd. Op die manier is 'wat' er geëxamineerd zou kunnen worden, beschreven.

In de kwalificatiedossiers is echter nergens beschreven 'hoe' er geëxamineerd moet worden, en welke inhoudelijke aandachtspunten daarbij gelden (bijvoorbeeld welke competenties per se voldoende moeten zijn of per se niet onvoldoende mogen zijn). Daarmee geven de dossiers veel ruimte voor interpretatieverschillen in de concrete uitwerking van (het niveau van) examens. Tevens is niet voorgeschreven 'waar' er geëxamineerd zou moeten worden en 'onder welke condities' (bijvoorbeeld: aantal

assessoren, kwaliteit van assessoren, individuele of groepsopdrachten, toegestane hulpmiddelen, e.d.).

De verantwoordelijkheid voor het aantonen van kwaliteit van de examens ligt bij de scholen zelf. Er is wel extern toezicht. Het KwaliteitsCentrum Examinering (KCE) was hiermee belast en had daartoe een serie standaarden beschreven waaraan examinering in het beroepsonderwijs zou moeten voldoen. KCE zag ook toe op de naleving van de standaarden middels audits. De taak van KCE is inmiddels overgenomen door de Inspectie van het Onderwijs. Een van de belangrijkste standaarden uit het Toezichtskader voor de examinering in het MBO is standaard 5:

'Het exameninstrumentarium voldoet aan toetstechnische kwaliteitseisen.'

In de toelichting op de standaarden van het Toezichtskader wordt de toetstechnische kwaliteitseis nader gespecificeerd met de termen validiteit en betrouwbaarheid. Een opleiding krijgt op dit aspect de beoordeling 'goed' als '....het examen dat gebruikt wordt aantoonbaar valide en betrouwbaar is' en de beoordeling 'voldoende' als het '....grotendeels valide en betrouwbaar is.'

Deze operationalisatie van de standaard is problematisch. In de eerste plaats omdat de kwaliteit wordt afgemeten aan onduidelijke criteria. Wat verstaat men precies onder betrouwbaarheid en validiteit⁴ en wat onder de termen 'aantoonbaar' en 'grotendeels'? In de tweede plaats omdat men aan instellingen om informatie vraagt die niet of slechts in zeer beperkte mate geleverd zal kunnen worden, aangezien het daarvoor noodzakelijke onderzoek niet alleen tijdrovend is maar ook specifieke deskundigheid vereist die vaak niet voorhanden is binnen de instelling.

Voor het vertrouwen van belanghebbenden in het MBO-diploma is juist deze standaard 5 van groot belang.

Leggen we standaard 5 naast de examinering van taal en rekenen/wiskunde dan doet zich nog een extra complicatie voor. Veel van de beroepsgerichte competenties worden geëxamineerd in de beroepspraktijk. Bij het vormgeven van toetsing van taal in de beroepspraktijk lijkt de deelvaardigheid 'gesprekken' nog het best een plek te krijgen bij geïntegreerde vormen van toetsing. Van beoordelaars in de praktijk kan echter niet zonder meer verwacht worden dat ze de "gespreksvaardigheid" kunnen beoordelen volgens de schalen van het CEF. Dit vergt nu eenmaal training en expertise op een vakgebied die de beoordelaars in de praktijk veelal niet hebben. Bij de huidige examinering zal, daar waar taalvaardigheid een plek heeft in de beoordeling in de beroepspraktijk, er dus waarschijnlijk variatie zijn in de beoordeling en zeker geen dekkende beoordeling zijn van alle deelvaardigheden (Lezen, Luisteren, Spreken/Gesprekken, Schrijven,). Deze praktijk, gekoppeld aan de eerder genoemde eisen voor het verkrijgen van vertrouwen in de examinering, zetten het vertrouwen in de examinering van taal ernstig onder druk.

Hoewel er nog geen ervaring is opgedaan met het *Raamwerk Rekenen-Wiskunde* gaat deze zelfde redenering op voor rekenen-wiskunde: bij een geïntegreerde beoordeling in

⁴ Voor beide begrippen kunnen in de literatuur over onderwijskundig meten veel verschillende operationalisaties worden aangetroffen.

de beroepspraktijk zal het niet mogelijk zijn om alle vereiste rekenvaardigheden aan bod te laten komen en om ze op de juiste manier te beoordelen.

In zijn algemeenheid geldt dat vertrouwen in examinering nauw samengaat met het eenduidig hanteren van een norm voor zakken en slagen. Als de kans groot is dat je met een zelfde resultaat op de ene school zakt en op de andere slaagt, leidt dat onherroepelijk tot wantrouwen in het examen. Eenduidige beslissingen over zakken en slagen zijn alleen mogelijk als het helder is wanneer een examenresultaat voldoende is en wanneer niet. Dat vraagt in de eerste plaats om een prestatiestandaard. Het hanteren van eenduidige beschrijvingskaders als het *Raamwerk MVT*, het *Raamwerk Nederlands* en het *Raamwerk Rekenen/wiskunde*, is daarbij een goede ontwikkeling. Maar dit zijn slechts beschrijvingskaders. De omschrijving van wat een kandidaat moet kennen en kunnen moet vervolgens 'vertaald' worden naar een kritische score op een of andere schaal waarop de examenprestatie kwantitatief wordt uitgedrukt.

Geconstateerd moet worden dat er met de overgang naar competentiegerichte gerichte beschrijvingen in de kwalificatiedossiers wel prestatiestandaarden beschreven zijn voor de beroepscompetenties maar niet of in mindere mate voor taal en rekenen/wiskunde. Dat is vermoedelijk het gevolg van de veronderstelling dat met het vaststellen van de bekwaamheid om de kerntaken van een beroep uit te voeren tevens is vastgesteld dat de daarvoor voorwaardelijke kennis en vaardigheden, waaronder die op het gebied van taal en rekenen/wiskunde, ook beheerst worden. Het ontbreken van heldere prestatiestandaarden impliceert uiteraard dat er ook geen valide kritische scores kunnen zijn om eenduidige beslissingen te nemen over de examenprestaties van kandidaten.

Voor taal is de oplossing gezocht in het beschrijven van niveaus van taalvaardigheid volgens het Common European Framework of Reference for Languages. Door een CEF-niveau aan te geven wordt verwezen naar de prestatiestandaarden in dat CEF (en in de speciaal voor het mbo ontwikkelde beschrijvingskaders *Raamwerk MVT* en *Raamwerk Nederlands*). Dit is echter nieuw en er zijn in het mbo nog niet veel ervaringen opgedaan met examens voor taal waarbij de prestatiestandaarden van het CEF worden toegepast⁵. Voor rekenen-wiskunde zal nog eerst ervaring opgedaan moeten worden met de prestatiestandaarden uit het *Raamwerk Rekenen/Wiskunde* dat pas op 13 december 2007 is uitgekomen.

⁵ Enkele ROC's werken al een tijdje met het Europees Taalportfolio en gebruiken daarbij de prestatiestandaarden van het CEF.

3 Advies voor een examineringsmodel

In dit hoofdstuk gaan we in op een model van examinering voor taal en rekenen/wiskunde. Allereerst geven we aan welke algemene uitgangspunten we hanteren bij dit advies.

3.1 Uitgangspunten

De volgende vijf uitgangspunten nemen we als basis voor ons advies:

1. We beogen een verhoging van de kwaliteit van examinering voor taal en rekenen/wiskunde.
2. We gaan uit van het *Common European Framework of Reference* (en de afgeleide documenten *Raamwerk MVT* en *Raamwerk Nederlands*) als basis voor het advies over examinering van taal.
3. We gaan uit van het *Raamwerk Rekenen/wiskunde voor het mbo (versie 0.9)* als basis voor het advies over examinering rekenen/wiskunde.
4. We realiseren ons dat diplomering in het beroepsonderwijs gericht is op een drievoudige kwalificering: voor het beroep, voor de doorstroom en voor de maatschappij. We gaan uit van de eisen voor talen en rekenen/wiskunde in de kwalificatiedossiers én in het document *Leren, Loopbaan en Burgerschap* (we gaan er van uit dat er een eis komt voor één mvt en voor rekenen/wiskunde in het document LL&B)
5. We onderschrijven het belang dat, gegeven de doelgroep en het beroepsgerichte karakter van opleidingen, examinering in het MBO zoveel mogelijk gekoppeld moet zijn aan functionele contexten.

Hieronder lichten we ieder uitgangspunt kort toe.

Ad 1.

De waarde van een diploma wordt bepaald door de kwaliteit van de examinering. Zoals bij elk meetinstrument wordt die kwaliteit beschreven in termen van betrouwbaarheid en validiteit. Betrouwbaarheid gaat over de meetnauwkeurigheid en validiteit over de geldigheid van de interpretaties die aan de examenresultaten gegeven worden. Betrouwbaarheid is een voorwaarde voor valide metingen en bij het ontwerpen van examens moet dan ook in de eerste plaats aandacht gegeven worden aan die aspecten die de betrouwbaarheid kunnen bevorderen. De belangrijkste aspecten betreffen het gebruik van kwalitatief hoogwaardige examenopgaven, het reduceren van de kans op 'toevalstreffers' (zoveel mogelijk opgaven), het terugdringen van subjectiviteit bij het beoordelen van de prestaties en de beschikbaarheid van een kritische score waartegen de examenprestaties kunnen worden afgezet.

Ad 2. Het Common European Framework of Reference

De eisen voor taalvaardigheid in het Nederlands en in de moderne vreemde talen (MVT) zijn in de kwalificatiedossiers vastgelegd in zogenaamde 'taalcompetentieprofielen'. Deze

taalcompetentieprofielen zijn gebaseerd op het Europees Referentiekader voor de talen (ERK). Het ERK is een transparante beschrijving van zes niveaus van taalvaardigheid, onderscheiden naar de vijf vaardigheden Luisteren, Lezen, Spreken, Gesprekken voeren en Schrijven. Het ERK is door de Raad van Europa ontwikkeld en door de Europese Unie geaccepteerd als algemene standaard voor het beschrijven van niveaus van taalvaardigheid bij het leren van een vreemde en tweede taal. Speciaal voor de moderne vreemde talen in het mbo is er een bewerking hiervan gemaakt: *het Raamwerk MVT*. Voor Nederlands is sinds december 2007 het '*Raamwerk Nederlands voor het (v)mbo*' beschikbaar. Dit was er echter nog niet toen de taalcompetentieprofielen voor Nederlands door de Kenniscentra werden opgesteld. Toen is gebruik gemaakt van het Raamwerk NT2, voor inburgeraars die Nederlands als tweede taal leren. In dit advies gaan we echter uit van het nieuwe beschrijvingskader: het *Raamwerk Nederlands*. In Bijlage 2 is een wat uitgebreidere beschrijving opgenomen van de taalniveaus, in het bijzonder van het drempelniveau B1.

Ad 3 Het Raamwerk rekenen-wiskunde in het mbo

Het *Raamwerk rekenen-wiskunde*, met beheersingsniveaus en deelgebieden is, net als het Raamwerk Nederlands, een instrument om nader aan te geven wat de deelnemers aan rekenen/wiskunde moeten beheersen. Het heeft herkenbare ijkpunten en is, anders dan een curriculumbeschrijving, gericht op het *functioneel gebruik* van rekenen/wiskunde. Het gaat bij rekenen/wiskunde om verschillende soorten vaardigheden met algemeen herkenbare niveaus, die in het gebruik altijd een specifieke kant hebben. Ze worden binnen het mbo ingezet in specifieke contexten: de beroepsgerichte contexten die te vinden zijn in de kwalificatiedossiers en de bredere maatschappelijke contexten die te vinden zijn in het document Leren, Loopbaan en Burgerschap. 'Wiskundige competentie' is een andere manier om aan te duiden dat het gaat om het vaardig en met inzicht kunnen en 'durven' inzetten van rekenen en wiskunde ten behoeve van het beroep, ten behoeve van burgerschap en ten behoeve van een verdere schoolloopbaan. Niet alleen de reken/wiskundige inhouden en vaardigheden, maar ook de complexiteit en vertrouwdheid van de context bepalen daarbij het niveau van handelen. Het raamwerk maakt het mogelijk om beheersingsniveaus binnen rekenen/wiskunde van elkaar te onderscheiden.

Ad 4 Drievoudige kwalificering

Het uitgangspunt van een mbo-diploma is dat het een drievoudige kwalificatie inhoudt: een kwalificatie voor het beroep, een kwalificatie voor de doorstroom en een kwalificatie voor de maatschappij. Naast de eisen voor taalvaardigheid⁶ voor Nederlands en de mvt's die zijn opgesteld vanuit de eisen van het beroep zijn er voor Nederlands (en binnenkort ook voor één mvt) ook niveau-eisen vastgesteld in het document 'Leren, loopbaan en Burgerschap' vanuit de notie dat taalvaardigheid een voorwaarde is om te kunnen leren op school, door te kunnen stromen naar een vervolgopleiding en als burger te kunnen functioneren in de maatschappij. Ook voor rekenvaardigheid zullen op een vergelijkbare manier niveau-eisen geformuleerd worden in het document LLB. Bij het nadenken over een model voor examinering van taal- en rekenvaardigheid geldt dat we hierbij redeneren vanuit alle drie de invalshoeken van het mbo-diploma.

⁶ De verschillende taalvaardigheden tezamen in een bepaalde context worden breed opgevat als 'taalcompetentie'.

Ad 5 Examinering in de context van het beroep

De doelgroep van het mbo is zeer divers. Over de gehele linie betreft het een doelgroep voor wie leren sterk gekoppeld is aan de toepasbaarheid van het geleerde in de context. Hoe meer en hoe verder geabstraheerd wordt van de concrete werkelijkheid, hoe moeilijker de mbo-deelnemer zal kunnen begrijpen waar het over gaat. Dit besef is van belang voor de onderwijsgeevenden. Ze zullen ervoor moeten zorgen dat de deelnemer dat wat hij moet leren, waar hij leert en hoe hij leert heel direct en concreet kan koppelen aan een beroepscontext. Ook leermiddelenmakers zullen zich hiervan bewust moeten zijn. Lesmateriaal moet herkenbaar zijn voor en toepasbaar zijn in de beroepscontext. Ook ontwikkelaars van examens hebben met dit gegeven te maken. Examinering moet zeer concreet gekoppeld zijn aan de beroeps- en/of leren, loopbaan en burgerschapscontexten die de mbo-er herkent.

3.2 In gebruik zijnde varianten van examinering

Examinering kan in meerdere of mindere mate gestandaardiseerd zijn. Standaardisatie kan de kwaliteit van een examen verhogen vanwege de toenemende controle op zowel het instrument als de procedure van de afname van het examen. Bovendien verhoogt standaardisatie de vergelijkbaarheid van de uitkomsten. Over het algemeen kan worden gesteld dat standaardisatie leidt tot een verhoging van de betrouwbaarheid van een examen.

Zoals gezegd kan standaardisatie betrekking hebben op de ontwikkeling van het exameninstrument zelf en op de afnameprocedure ervan. Als zowel de ontwikkeling als de afname van het examen door regelgeving volledig gestandaardiseerd verlopen, heeft een examen het volgende karakter: alle kandidaten maken op een zelfde moment dezelfde examenopgaven onder voorgeschreven examencondities. De antwoorden worden (machinaal) gescoord en de beslissing 'voldoende/onvoldoende' wordt gebaseerd op een voorgeschreven cesuurscore. Zo'n aanpak leidt tot de meest objectieve en daardoor vergelijkbare examenresultaten.

De schaduwzijde van een dergelijke aanpak is dat ze vaak leidt tot een drastische abstrahering van wat werkelijk beheerst en dus gemeten zou moeten worden. Vraag is dan dus of we nog wel meten wat we wilden meten. Oftewel, of de validiteit van het examen met een dergelijke examenvorm te borgen is. Bij het kiezen van een model voor examinering is het zaak een goede afweging te maken tussen de betrouwbaarheid van de meting aan de ene kant en validiteit van de meting aan de andere kant. Afhankelijk van de te nemen beslissing, de gewenste controle op de te nemen beslissing en de mogelijkheden voor uitvoering ervan (in praktisch en/of financieel opzicht) kan in meer of mindere mate gestreefd worden naar standaardisatie van de examinering.

Examens kunnen op verschillende manieren al dan niet gestandaardiseerd zijn. Hieronder worden een drietal in gebruik zijnde varianten besproken die het spectrum van standaardisatie van examinering dekken: van meer naar minder toezicht op examenproductontwikkeling en examenafname.

3.2.1 Variant 1 Inburgeringsvariant

In deze variant is er toezicht op zowel de ontwikkeling van het product als op de afname ervan in de praktijk. Onder directe verantwoordelijkheid van de overheid wordt de ontwikkeling van een examen uitgevoerd. Daarmee is er rechtstreeks toezicht op de kwaliteit van het examenproduct. De afname van het examen vindt in deze variant eveneens plaats onder direct toezicht van de overheid. Op speciale, door de minister aan te wijzen, examenlocaties vindt de examinering plaats. Dit gebeurt onafhankelijk van de onderwijsgeveden.

3.2.2 Variant 2 Voortgezet onderwijs variant

Ook in deze variant wordt het examen centraal ontwikkeld. Daarmee is er regie op de inhoud en kwaliteit van het examen zodat betrouwbaarheid en validiteit geborgd kunnen worden. Het examen wordt echter decentraal afgenomen. Iedere onderwijsaanbieder draagt zelf zorg voor een correcte uitvoering van de afname. De verwerking van de antwoorden gebeurt echter weer centraal. Middels allerlei procedures en controles kan toezicht gehouden worden op de uitvoering. Hierin is variatie mogelijk.

3.2.3 Variant 3 MBO-variant

In de MBO-variant is de onderwijsaanbieder zelf verantwoordelijk voor de borging van de kwaliteit van zowel het examen als de afname. Daarmee vinden examenontwikkeling en examenuitvoering decentraal plaats. Ook in deze variant is het mogelijk meer of minder controle in te bouwen op het toezicht op zowel ontwikkeling van uitvoering van de examinering.

3.2.4 Combinaties van varianten

Combinaties van deze varianten zijn mogelijk en komen voor. Het examen voor het voortgezet onderwijs is feitelijk ook een mengvorm van 2 en 3 omdat het uiteindelijke eindcijfer voor de helft bepaald wordt door het centraal ontwikkeld examen. De andere helft van het cijfer wordt immers bepaald door een schoolexamen. Ook het inburgeringsexamen is feitelijk een combinatie van variant 1 en 2. Ook daar is een combinatie van centraal ontwikkelde examenonderdelen die centraal worden afgenomen en centraal ontwikkelde delen die decentraal worden afgenomen. Als de ontwikkeling van examenproducten en het afnameproces niet onder directe regie van de overheid plaatsvindt, is meer controle nodig op de decentrale borging van kwaliteit van examenproduct en proces. In de eerste variant hoeft er decentraal vrijwel niets geregeld of gedaan te worden. In de derde variant is de uitwerking van de controle uitermate belangrijk. Alleen dan kan een examen met een civiel effect worden geborgd.

3.2.5 Welke variant is het meest geschikt voor taal en rekenen/wiskunde in het mbo?

Varianten zijn niet goed of fout. Bij het kiezen en vervolgens concreet vormgeven van een model voor examinering komt het erop aan te komen tot de best passende variant. Het best passend waar het gaat om de afweging tussen de betrouwbaarheid, de validiteit en

de praktische uitvoerbaarheid. Maar ook moet de vorm van examinering passen bij de vorm van het competentiegerichte onderwijs in het mbo. De manier van examinering stuurt immers sterk de manier waarop het onderwijs wordt vormgegeven. Dit sturende effect moet worden meegewogen bij een keuze voor een examineringsmodel. Daarnaast zijn uiteraard ook financiële aspecten van belang.

Afhankelijk van de gekozen variant moeten aanvullende procedures voor controle worden uitgewerkt om te komen tot een examinering die herkend en erkend wordt door alle betrokkenen.

De volgende, meer concrete, vragen zijn belangrijk om tot een keus te kunnen komen:

- Wat moeten we meten? Welke beslissing(en) moeten we nemen?
- Hoe belangrijk is de beslissing die we nemen over de kandidaten?
- Hoe groot is de kans op fraude? Hoe aantrekkelijk is frauderen?
- In hoeverre zijn de met het examen te nemen beslissingen plaatselijk of regionaal gebonden?
- Kun je met een gestandaardiseerd examen de beslissing nemen die je met het examen wilt nemen?
- In hoeverre zijn (vak)deskundige beoordelaars nodig voor de beoordeling?
- In hoeverre is de expertise aanwezig voor decentrale of centrale examenontwikkeling en uitvoering?
- In hoeverre is het examen decentraal af te nemen? Welke faciliteiten zijn daarvoor nodig? Welke zijn daarvoor beschikbaar?
- Welke financiële middelen zijn gemoeid met de keuze voor een examenmodel?
- Hoe verhouden de kosten zich tot de te nemen beslissingen?

De MBO raad en COLO hebben al eerder aangegeven te willen koersen op meer standaardisering van het gehele mbo-examen. Er wordt in dit kader gesproken van 'examenprofielen' per kwalificatiedossier: een soort format voor een examen. Dit lijkt ons op het eerste gezicht een goede ontwikkeling. Ook wordt er gesproken over de mogelijkheid om voor de onderliggende kennis en vaardigheden gestandaardiseerde toetsen te ontwikkelen en af te nemen. Hierbij wordt nadrukkelijk gedacht aan de talen en rekenen/wiskunde.

Voor taal wordt vaak het argument aangevoerd dat lokale en/of regionale behoeften nauwelijks een rol spelen. Ook is het niveau voor taal op de werkplek niet goed te beoordelen. Hoe kan een werkgever vaststellen of een kandidaat daadwerkelijk een tekst kan lezen op taalniveau B2? Daar staat tegenover dat ervaring leert dat kandidaten bepaalde algemene vaardigheden weer het beste kunnen demonstreren in een voor hen relevante context. Een kandidaat die moet aantonen dat hij gesprekken kan voeren op taalniveau B1 zal dat naar alle waarschijnlijkheid het beste kunnen laten zien in een gesprek met een zorgvrager als het gaat om een kandidaat die de opleiding Verzorgende volgt. In die zin is taal méér dan de losse, contextloze vaardigheden en moet taal, conform de uitgangspunten van het CEF, gezien worden als een brede competentie.

Een vergelijkbare problematiek komt naar voren bij rekenen/wiskunde, hoewel de verschijningsvorm anders is. Voor de meeste reken/wiskundige vaardigheden maakt het

een groot verschil of de vaardigheden moeten worden getoond binnen een voorstelbare (beroeps)context of op zichzelf staand en dus zonder context. In het eerste geval is er altijd sprake van betekenisvol handelen waarbij de context de betekenis geeft aan het reken/wiskundig handelen, terwijl in het tweede geval vaardigheden worden gevraagd op een hoger abstract niveau zonder binding met enige (beroeps)context.

Voor rekenen/wiskunde is er op dit gebied alleen ervaring binnen de sector Techniek (TWIN-project). Daar is gebleken dat wiskundige kennis en vaardigheid bij veel deelnemers voldoende aanwezig blijkt te zijn als de probleemstelling direct is gekoppeld aan een voorstelbare probleemsituatie die is gerelateerd aan de beroepssector, maar dat een vergelijkbare bevraging op een hoger abstractieniveau, los van een specifieke (beroeps)context, zoals aan de orde in het doorstroomprogramma dat voorbereide op een voortzetting van de studie op HBO-niveau, onoverkomelijke problemen kan opleveren.

Deze ervaringen sporen heel goed met bevindingen in Engeland, waar is geconstateerd dat beroepsbeoefenaren bij het oplossen van problemen wiskundige vaardigheden altijd verweven met specifieke methodieken binnen de beroepscontext. De Engelse onderzoekers spreken in dit verband van *situated abstraction*, waarbij de beroepsbeoefenaren *anchors* uit de beroepscontext gebruiken (Hoyle&Noss, 1998). Wanneer die anchors worden weggehaald en vergelijkbare problemen, contextvrij (abstract) of in een andere context (transfer), moeten worden opgelost, blijken de meesten van hen (bankmedewerkers, verpleegkundigen en piloten) volledig vast te lopen. Op zulke momenten lijken de binnen de beroepscontext zelfverzekerde personen terug te moeten vallen op methodieken van de schoolwiskunde en dat maakt ze zichtbaar nerveus en onbekwaam.

Bij een opvatting over rekenen/wiskunde als een brede competentie lijkt het voor rekenen/wiskunde mogelijk om de beoordeling van de kennis en vaardigheden binnen de beroepscontext te laten plaatsvinden mits er voldoende duidelijkheid wordt gegeven in de vorm van standaarden voor toetsing en beoordeling. Dit is een belangrijk aandachtspunt bij de verdere uitwerking van het raamwerk.

Daarnaast vereist examinering gericht op doorstroom naar een hoger niveau van opleiden (waarbij eerder is gesteld dat een verhoging van onderwijsniveau automatisch een hoger abstractieniveau met zich meebrengt) een hoger beheersingsniveau binnen het raamwerk met zich meebrengt. Voor de doorstroom van mbo (niveau 4) naar hbo⁷ geldt daarbij ook dat het onderwijs van het mbo wordt gekarakteriseerd als *inductief*, terwijl het hbo wordt gekenmerkt door het *deductieve* karakter van de opleiding.

Een examenmodel voor taal en rekenen/wiskunde moet zo goed mogelijk aansluiten bij de hier geschetste feiten en ervaringen en bij de uitgangspunten van competentiegericht beroepsonderwijs.

⁷ LICA (september 2006) *Onderzoeksverslag Duurzaamheid en versterking doorstroom mbo-hbo 'Pontons tussen twee bewegende oevers'* p 17

3.3 Advies voor examinering van taal- en rekenvaardigheid

3.3.1 *Het examenmodel algemeen*

Competentiegericht opleiden (CGO) vraagt om een vorm van examinering waarin kandidaten moeten laten zien dat ze beschikken over de competenties die vereist zijn voor de uitvoering van authentieke beroepstaken. Problematisch in dit verband is dat dergelijke praktijkexamens niet of nauwelijks te standaardiseren zijn waardoor niet alleen de kwaliteit van deze examens maar ook het niveau van de gediplomeerden niet volledig gegarandeerd kan worden. Het gevolg daarvan kan zijn dat, bijvoorbeeld, een bedrijfsadministrateur niveau 4 uit Amsterdam niet zonder meer uitwisselbaar is met een uit Maastricht.

De vraag is hoe er recht gedaan kan worden aan de eisen die CGO stelt aan examinering met behoud van de vergelijkbaarheid van diploma's. Of anders gezegd hoe het mbo-examen zó ingericht kan worden dat de uitslagen betrouwbaar en valide te interpreteren zijn. Helaas staan betrouwbaarheid en validiteit in een complexe verhouding tot elkaar. Maatregelen die de betrouwbaarheid kunnen opvoeren, zoals standaardisatie van de inhoud en/of van afnameprocedures, pakken vaak ongunstig uit voor de representativiteit van de exameninhouden en dus voor de validiteit van het examen en andersom zijn natuurgetrouwe examenopdrachten niet goed te standaardiseren waardoor problemen ontstaan met de objectiviteit en meetnauwkeurigheid van de scores.

Een voor de hand liggende oplossing is de toepassing van een hybride examenmodel, wat inhoudt dat het examen zal bestaan uit een mix van decentraal en centraal ontwikkelde examens.

Decentraal ontwikkelde examinering

Of een kandidaat de competenties verworven heeft om de kerntaken van een uitstroomprofiel adequaat uit te voeren is in principe alleen vast te stellen door het functioneren van de betreffende persoon in verschillende (zo authentiek mogelijke) taaksituaties te beoordelen om zo een bewijsdossier op te bouwen waarmee gerechtvaardigde beslissingen over de beroepsbekwaamheid te nemen zijn. Centraal ontwikkelde examens voor beroepsmatig handelen zijn hier niet goed denkbaar vanwege het feit dat:

- opdrachten vaak regionaal 'gekleurd' en soms zelfs bedrijfsspecifiek worden ingevuld;
- deelnemers steeds meer mogelijkheden krijgen om het onderwijsleerproces te individualiseren waardoor 'vaste' examenmomenten steeds meer plaats maken voor flexibele examenmomenten. Geheimhouding van examenopdrachten zou vragen om een enorme hoeveelheid gelijkwaardige examenopdrachten voor elk uitstroomprofiel.

De onderlinge vergelijkbaarheid van dergelijke bewijsdossiers vraagt grote inspanningen voor borging doordat opdrachten, contexten waarbinnen de opdrachten worden

uitgevoerd, assessoren en ingezette assessmentvormen (hands-off, simulaties, hands-on) van kandidaat tot kandidaat kunnen verschillen. Kwaliteitsborging van het examen kan praktisch gezien alleen plaatsvinden door controle op naleving van de voorschriften ten aanzien van de vulling en de beoordeling van de bewijsdossiers.

Centraal ontwikkelde examinering

Centraal ontwikkelde examens zijn een geschikte examenvorm als het mogelijk is om grote groepen kandidaten op hetzelfde moment aan dezelfde opgaven te laten werken. Voor de talen en rekenen/wiskunde is dit, in ieder geval op onderdelen, het geval. Beheersing van lees- en luistervaardigheid leent zich prima voor examinering door middel van gestandaardiseerde toetsen. Echter, de meer productieve taalvaardigheden zoals spreken/gesprekken voeren en schrijven zijn te bewerkelijk voor centraal ontwikkelde examinering.

Bij rekenen/wiskunde is een centraal ontwikkeld examen haalbaar om vast te stellen of de kandidaten beschikken over de reken-/wiskundige gereedschappen die nodig zijn voor een adequate uitvoering van beroepstaken. Maar de bekwaamheid om die gereedschappen in authentieke taaksituaties in te zetten leent zich weer veel minder voor centrale examinering aangezien de hogere cognitieve vaardigheden (zoals probleemoplossen) die daarbij in het geding zijn, niet goed in beeld gebracht kunnen worden met gestandaardiseerde, schriftelijke (lieft geautomatiseerde) toetsvormen.

Een hybride model

Om de genoemde redenen pleiten wij voor een hybride model van examinering in het mbo, waarin een evenwicht gezocht wordt tussen de verschillende voor- en nadelen en waarin de dilemma's zoveel mogelijk met elkaar worden verzoend. In dit hybride model is plaats voor zowel centraal ontwikkelde, gestandaardiseerde examens als decentraal ontwikkelde onderdelen. Een dergelijk hybride model zou ook toegepast kunnen worden voor taal en rekenen/wiskunde.

Belangrijk bij de ontwikkeling van standaard examens (als onderdeel van het hybride model) is dat er wordt gewerkt met een en dezelfde norm. We sluiten daartoe aan bij het CEF (en de Raamwerken MVT en Nederlands) en het Raamwerk Rekenen/wiskunde. De kwalitatieve prestatiestandaarden in deze beschrijvingskaders worden omgezet in kwantitatieve toetsnormen of cesuren. Per te onderscheiden deelvaardigheid⁸ gaan we uit van één en dezelfde meetlat waarop we verschillende cesuren aanbrengen. Vergelijk dit met een thermometer. We ontwikkelen per deelvaardigheid een thermometer. Bij een bepaald aantal graden Celsius geven we aan dat een kandidaat het algemene basisniveau, bijvoorbeeld lezen A2, beheerst. Bij een hoger aantal graden Celsius geven we aan dat de kandidaat het niveau beheerst dat volgens het kwalificatiedossier noodzakelijk is om in het beroep te kunnen starten op de arbeidsmarkt, bijvoorbeeld Luisteren B1. Bij een nog hoger aantal graden Celsius geven we aan dat een kandidaat kan doorstromen naar een HBO-opleiding, mogelijk gekoppeld aan een bepaalde richting (bijvoorbeeld voor bepaalde reken/wiskunde vaardigheden). We streven er ook naar om

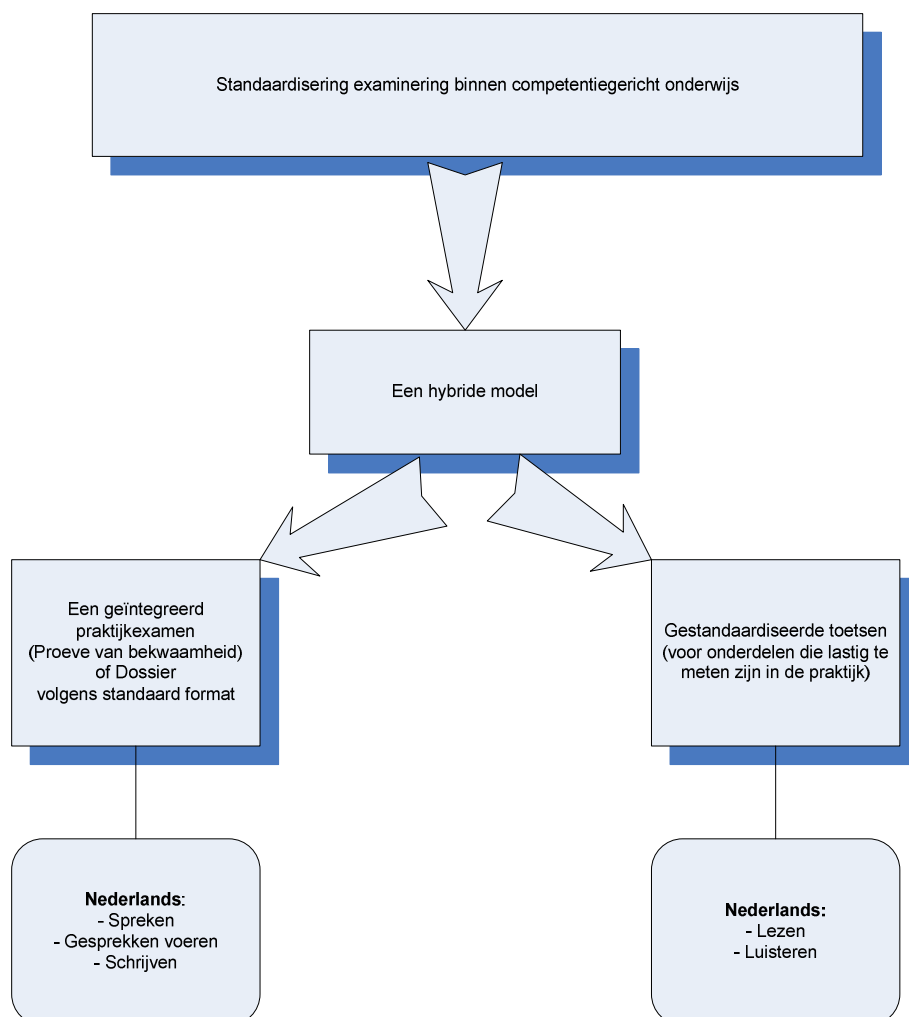
⁸ Hiermee doelen we bij taalvaardigheid op de door het CEF onderscheiden deelvaardigheden Spreken/Gesprekken, Luisteren, Lezen en Schrijven. Bij rekenvaardigheid gaan we uit van de deelgebieden Getallen/hoeveelheden en maten, Ruimte en vorm, Gegevensverwerking en onzekerheid, Verbanden en veranderingen

voor doorstroom van de niveau 1-2 opleidingen naar een niveau 3 opleiding een 'temperatuurniveau' aan te geven op diezelfde thermometer.

In de volgende paragrafen wordt het voorgestelde hybride examenmodel nader uitgewerkt voor Nederlands, een moderne vreemde taal en rekenen/wiskunde.

3.3.2 **Examenmodel voor Nederlands**

We adviseren de invoer van een hybride examenvariant voor Nederlands. Met hybride bedoelen we dat er een gestandaardiseerd deel centraal wordt ontwikkeld en een deel wordt overgelaten aan de school zelf. Echter, ook voor dit laatste deel adviseren we een vorm van standaardisering in de vorm van een format of 'examenprofiel'.



We adviseren om centraal een gestandaardiseerd exameninstrument voor de Lezen en Luisteren te ontwikkelen. Dit examen moet echter decentraal kunnen worden afgenomen (op de school zelf), op flexibele momenten. Dit vraagt om een geautomatiseerd instrument en een gecontroleerde afname op school. In hoofdstuk 4 wordt hier verder op ingegaan.

De keuze voor Lezen en Luisteren is ingegeven door de volgende argumenten:

- Deze zogenaamde 'receptieve' vaardigheden zijn moeilijk meetbaar in de beroepspraktijk
- Deze vaardigheden lenen zich goed voor een geautomatiseerd examen
- Door zowel voor Lezen als Luisteren te kiezen borgen we zowel een vaardigheid uit het mondelinge als uit het schriftelijke domein.

De examens voor Lezen en Luisteren werken met één meetlat, gerelateerd aan het CEF. We stellen voor echter examens aan te bieden in verschillende niveau-varianten die te herleiden zijn tot dezelfde meetlat zodat vergelijking mogelijk is (zie ook hoofdstuk 4 en bijlage 4). Een kandidaat kan bijvoorbeeld opteren voor het afleggen van een examen op niveau A2/B1 of een examen B1/B2, of B2/C1. Zo blijft de examenlast beperkt en is een nauwkeurige meting wel mogelijk. De ondergrens van het examen stellen we voor te leggen op het niveau van inburgering: taalniveau A2. Voor Nederlands in het mbo is het niet nodig een examen te ontwikkelen op niveau A1.

De taaltaken/taalhandelingen voor Lezen en Luisteren kunnen zich afspelen binnen brede contexten, zoals ook gebeurt bij de Staatsexamens Nederlands als Tweede Taal⁹. Dit is te rechtvaardigen door het feit dat Nederlands altijd nodig is en zich afspeelt binnen alle contexten. Het examen mag echter géén 'Voortgezet Onderwijs examen' worden met algemene contexten. De contexten zullen algemeen beroepsgericht moeten zijn én gericht op de kerntaken uit Leren, Loopbaan en Burgerschap.

Spreken/Gesprekken voeren en Schrijven kunnen door de school worden ontwikkeld en beoordeeld. Dit is ingegeven door de volgende argumenten:

- Deze zogenaamde 'productieve' vaardigheden komen in bijna alle kerntaken voor en kunnen goed worden aangetoond in de beroepspraktijk
- Deze vaardigheden lenen zich niet zo goed voor een geautomatiseerd examen
- Door voor zowel Spreken/Gesprekken voeren als Schrijven de mogelijkheid te bieden dit in de beroepspraktijk aan te tonen borgen we zowel een vaardigheid uit het mondelinge als uit het schriftelijke taalvaardigheidsdomein.

Ook de schoolexamens voor Spreken/Gesprekken en Schrijven werken met één meetlat, gerelateerd aan het CEF. Omdat het voor de meeste scholen nogal lastig is om dit zelf vorm te geven (en er daardoor verschillen kunnen ontstaan en dus de betrouwbaarheid in het geding is) stellen we voor de normen voor beoordeling van Spreken/Gesprekken voeren en Schrijven centraal te ontwikkelen. De uitvoering van de examinering kan decentraal plaatsvinden.

We stellen voor om een aantal instrumenten te ontwikkelen die de standaardisatie in de praktijk van de decentrale uitvoering zullen vergroten. Belangrijk is dat er beschrijvingen komen van taalhandelingen/taaltaken die een goede uitwerking vormen van het meten van Spreken/Gesprekken voeren en Schrijven. Ook het aantal bewijzen dat nodig is om aan te tonen dat iemand voldoende taalvaardig is wat betreft deze vaardigheden moet van te voren worden uitgewerkt.

⁹ Het Staatsexamen Programma I en II worden sinds 1992 in opdracht van het ministerie van OC&W ontwikkeld door Bureau ICE en Cito.

Naast het ontwikkelen van beschrijvingen van taalhandelingen/taaltaken adviseren we om de volgende twee instrumenten te ontwikkelen:

- een taalportfolio Nederlands¹⁰ waarin bewijzen verzameld kunnen worden van de vaardigheden Spreken/Gesprekken voeren en Schrijven, gekoppeld aan een uitgewerkte set criteria voor de verschillende niveaus.
- uitgewerkte beoordelingscriteria die opgenomen kunnen worden in praktijktoetsen, zoals proeven van bekwaamheid, zodat standaardisatie van de beoordelingscriteria wordt geborgd. Hieraan voegen we dan ook richtlijnen toe voor opdrachten die (nog) wel of niet (meer) geschikt zijn voor een bepaald niveau.¹¹

Voor Spreken/Gesprekken en Schrijven moeten de taaltaken/taalhandelingen die gemeten worden afgestemd worden op de beroepscontext van een specifieke kwalificatie. Ook de contexten van LLB kunnen hierbij worden betrokken. Het is de keuze van de school om de beoordeling vorm te geven binnen een proeve van bekwaamheid of met behulp van een Dossier van een taalportfolio of een combinatie daarvan. Het format moet handvatten bieden voor wat minimaal aan bewijslast vereist is. De controle op de afname van dit schoolexamen Nederlands voor Spreken/Gesprekken en Schrijven zou aan dezelfde standaarden moeten worden onderworpen als die gelden voor de examinering van beroepsgerichte competenties. Dit betekent onder andere dat de beoordeling moet gebeuren door gekwalificeerde assessoren.

Wij menen dat met dit hybride model een betrouwbare en valide toetsing van Nederlands in het mbo mogelijk is die tegelijkertijd praktisch uitvoerbaar is. Bovendien zal, door de hybride vorm, het sturend effect tweeledig zijn: enerzijds zal het onderwijs zich meer richten op expliciete aandacht voor het leren Lezen en Luisteren, anderzijds zal de aandacht uitgaan naar het vergroten van de taalvaardigheid binnen de context van het beroep.

3.3.3 Examenmodel voor de moderne vreemde talen (mvt)

Voor de examinering van de moderne vreemde talen stellen we in principe eenzelfde hybride model voor. Er zijn echter enkele wezenlijke verschillen met Nederlands:

- de mvt's worden in principe geleerd voor het beroep. Er wordt vooral vaktaal geleerd. Iemand die Duits leert voor Toerisme beheerst een andere woordenschat dan iemand die Duits leert voor Autotechniek.
- Er is (nog) geen verplichting voor een mvt in kader van Leren, Loopbaan en Burgerschap. Het advies over één mvt spreekt zich niet uit over een bepaalde taal, maar laat de keuze tussen Engels, Duits, Frans en Spaans. Ook is geen sprake van een minimaal niveau.

¹⁰ Binnen het Aanvalsplan Laaggeletterdheid wordt in 2008 een taalportfolio Nederlands ontwikkeld. De ontwikkeling hiervan zou moeten worden afgestemd op de keuze voor dit examenmodel.

¹¹ Zoals bijvoorbeeld het voor NT2 ontwikkelde instrument 'Stalenboek Assessment'

- Het aantonen in de beroepscontext is nóg lastiger dan voor het Nederlands (behalve bij een buitenlandse stage)
- Naast de veel voorkomende mvt's als Engels, Duits, Frans en Spaans bieden sommige scholen ook andere talen aan zoals Italiaans, Portugees of Turks. Deelnemers kunnen deze talen kiezen wanneer in het kwalificatiedossier de keuze voor de mvt vrij is.
- Veel leerlingen beginnen aan een vreemde taal op nul-niveau.

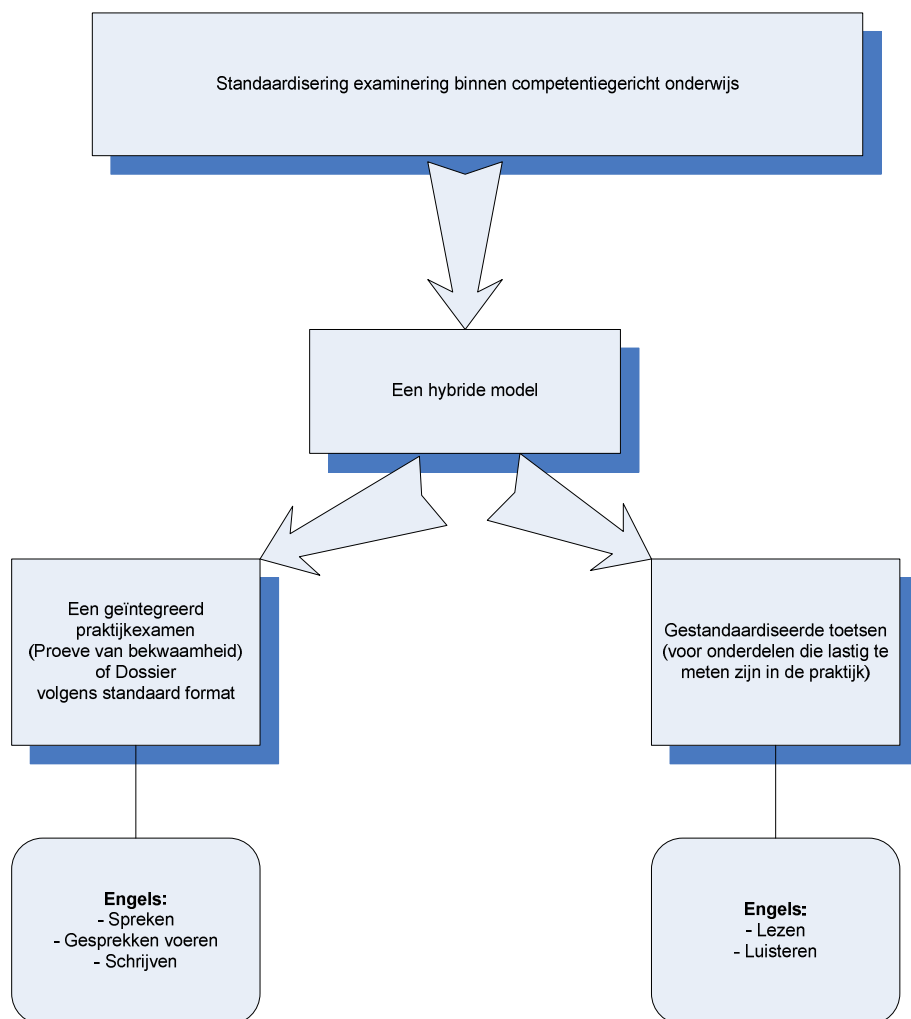
Deze andere positie van de mvt's maakt de ontwikkeling van een gestandaardiseerd examen voor Lezen en Luisteren complexer:

- Er zouden examens moeten worden ontwikkeld voor ten minste vier mvt's, maar wellicht voor méér.
- De examens zouden meer specifiek beroepsgericht moeten zijn. Te denken valt aan bijvoorbeeld varianten voor de verschillende sectoren (Techniek, Zorg, Economie). Op de lage niveaus (A1 en A2) zijn ook algemene contexten gerechtvaardigd.
- Er moeten examens komen waarin het A1 gebied ook getoetst kan worden. Echter, C1 kan buiten beschouwing worden gelaten omdat dit niet voorkomt in de kwalificatieprofielen.

Voor de ander pijler in het hybride model (Spreken/Gesprekken en Schrijven binnen de beroepscontext) is er het voordeel dat er al een Taalportfolio bve (inclusief een digitale variant) is voor de mvt's waar ook al ervaring mee is opgedaan. Bij de evaluatie van dit instrument is echter wel gebleken dat er grote behoefte is aan een nieuwe versie met verbeteringen naar aanleiding van deze ervaringen. Wij adviseren dan ook om, parallel aan de ontwikkeling van een nieuw Taalportfolio Nederlands een nieuwe versie te ontwikkelen van de Taalportfolio MVT.

Verder geldt, net als bij Nederlands, dat er formats en standaarden moeten worden ontwikkeld om de betrouwbaarheid van het schoolexamen te verhogen. Ook zullen er voor de beoordeling van de mvt's specifieke taalassessoren nodig zijn.

Gezien de complexiteit van de positie van de mvt's adviseren wij om alleen voor Engels gestandaardiseerde examens te ontwikkelen. Deze kunnen algemene maatschappelijke en algemeen beroepsgerichte contexten bevatten, net als bij Nederlands.



3.3.4 **Examenmodel voor Rekenen/wiskunde**

Het lijkt nog te vroeg om een definitief examenmodel voor rekenen/wiskunde voor te stellen. Een nadere uitwerking van het raamwerk zal duidelijkheid moeten verschaffen. Daarom volgt hier een beschouwing van mogelijke opties.

De vier deelgebieden uit het Raamwerk Rekenen/wiskunde, te weten

- Getallen/hoeveelheden en maten
- Ruimte en vorm
- Gegevensverwerking en onzekerheid
- Verbanden en veranderingen

zullen binnen de kwalificaties geheel verschillend worden gehanteerd. Afhankelijk van het beroepsprofiel zullen er verschillende niveau-eisen worden gesteld. De sector Techniek stelt immers geheel andere eisen op het gebied van wiskundige vaardigheden dan de Zorgsector. Zelfs binnen een sector zullen grote verschillen optreden. Het lijkt dan ook weinig zinvol om centrale examinering voor te stellen voor de eisen aan

Rekenen/Wiskunde die straks zullen worden opgesteld vanuit de eisen van het beroep (per kwalificatiedossier), juist omdat hier de beroepsgerichte context zo divers is.

Dat ligt anders bij de niveau-eisen voor Rekenen/Wiskunde in het kader van Burgerschap Voor Burgerschap is het goed voorstelbaar dat er een gestandaardiseerd examen komt voor het hele mbo. Het Raamwerk Rekenen-Wiskunde geeft een advies voor te hanteren uitstroomniveaus per mbo-niveau. De contexten van de kerntaken voor Burgerschap bieden voldoende aanknopingspunten voor de ontwikkeling van contextrijke examenopgaven.

Voor de doorstroom (van niveau 4 naar het hbo) is een sectorbrede gestandaardiseerde examinering, zoals in het verleden gehanteerd bij de sector Techniek, een mogelijke optie. De eisen op dit gebied zijn vergelijkbaar met de eisen die vanuit de verschillende profielen van havo worden gehanteerd. Deze examens werden ontwikkeld onder het wakend oog van de LEC wis- en natuurkunde mto en door Cito ontwikkeld en geëvalueerd. Dat is vergelijkbaar met de bij het VO gehanteerde constructie CEVO-Cito. Een dergelijke sectorspecifieke examinering zou daarom goed kunnen worden gekoppeld aan de CEVO. Er wordt vanuit CEVO ook gewerkt aan flexibilisering waardoor op dit moment er drie examenmomenten zijn: januari, mei en september.

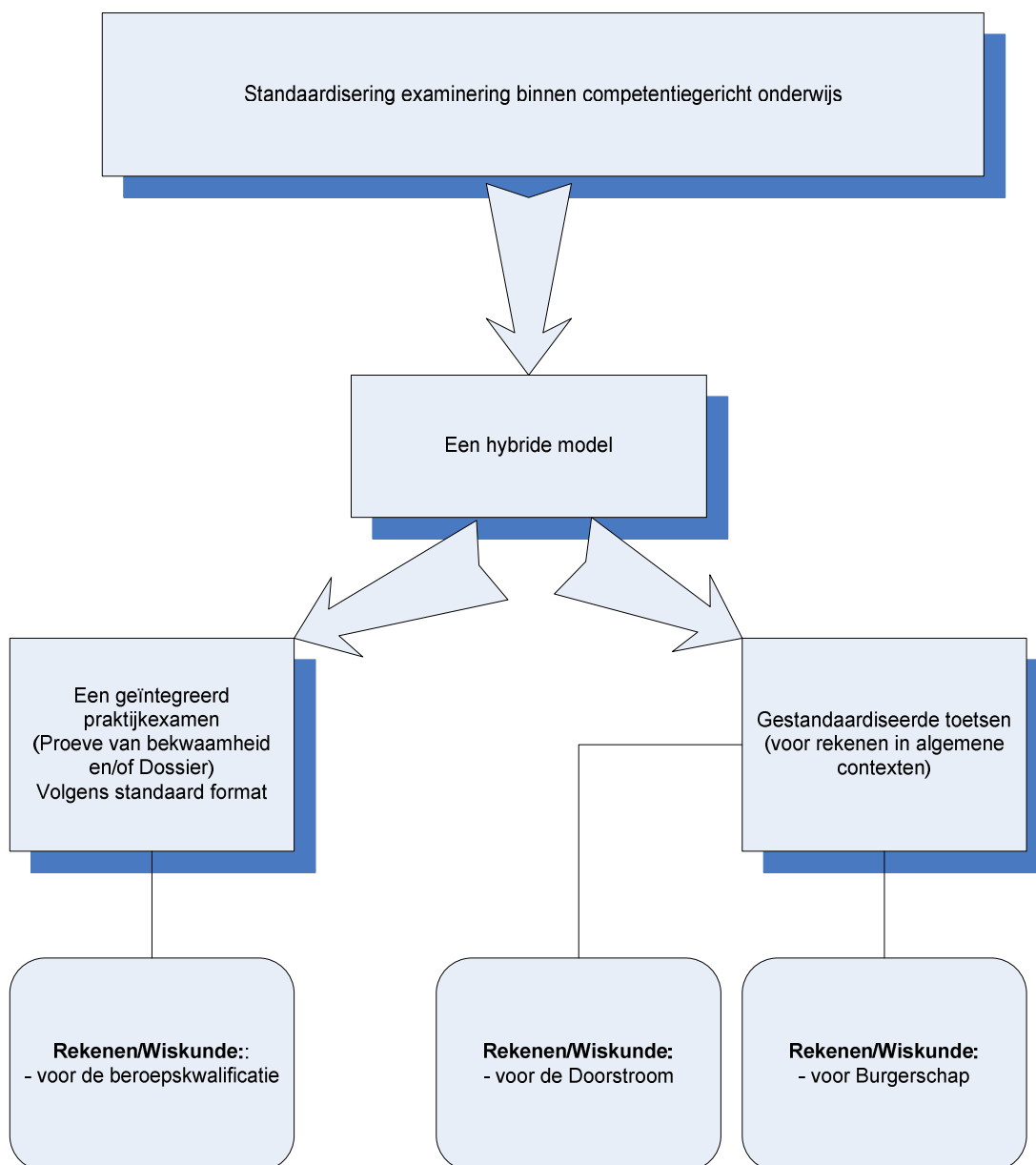
Hoewel gestandaardiseerde examinering voor Rekenen/Wiskunde binnen de beroepskwalificatie niet een voorstelbare constructie lijkt, adviseren we toch een zekere mate van standaardisering van het door te school af te nemen examen. Het is mogelijk om aan beroepstaken kwaliteitseisen te koppelen die moeten garanderen dat de vaardigheden op het gebied van rekenen/wiskunde binnen de beroepscontext kunnen worden beoordeeld. Dat vereist een beschrijving van standaarden en richtlijnen, een soort format¹².

Samenvattend adviseren we voor Rekenen/wiskunde het volgende:

- Ontwikkel een gestandaardiseerd examen voor de eisen voor Rekenen/wiskunde binnen de Burgerschapscompetenties. Een dergelijk examen moet flexibel en decentraal kunnen worden afgenomen.
- Doorstroomcompetenties voor Rekenen/wiskunde moeten per sector worden vastgesteld. Vervolgens kunnen ook hiervoor gestandaardiseerde examens worden ontwikkeld. De examens moeten flexibel en decentraal kunnen worden afgenomen.¹³
- De eisen voor Rekenen/Wiskunde binnen de beroepskwalificatie worden decentraal geëxamineerd (op school), zoveel mogelijk geïntegreerd met het beroepsmatig handelen. Hiervoor worden centraal standaarden en beoordelingscriteria ontwikkeld, waarmee controle op de kwaliteit kan worden uitgevoerd.

¹² Het ligt in de bedoeling om voorbeelden van dergelijke standaarden gekoppeld aan het raamwerk te ontwikkelen voor eind 2008, als de definitieve 1.0 versie van het Raamwerk zal worden opgeleverd.

¹³ De aansturing en bewaking kan mogelijk worden ondergebracht bij CEVO of een in te stellen orgaan van de MBO Raad (zoals de vroegere LEC-constructie)



3.4 Slagen of zakken?

Het kwalificatiedossier doet geen uitspraken over hoe scholen om moeten gaan met de beoordelingsresultaten op de verschillende onderdelen. Er wordt van uitgegaan van een holistische benadering: wanneer het geheel aan samenhangende competenties voldoende is kan de deelnemer worden gediplomeerd. Dit kan leiden tot niet vergelijkbare beoordelingen en dus diplomering.

Wanneer gekozen wordt voor het voorgestelde examenmodel voor de talen en rekenen/wiskunde met daarbinnen een gestandaardiseerde toets/toetsen komen we voor dezelfde vraag te staan: bij welke uitslag van het examen is een deelnemer geslaagd? Het ligt voor de hand dat het uitgangspunt hierbij zou moeten zijn dat een deelnemer het vereiste niveau voor de talen (alle vaardigheden) én rekenen/wiskunde behaald moet

hebben om het mbo-diploma met de drievoudige kwalificatie te krijgen. Maar dan moet nog nagedacht worden over wanneer iemand een voldoende heeft voor bijvoorbeeld taal. Moeten dan alle onderdelen voldoende zijn? Zijn onderdelen onderling compensabel? Hoe verhoudt het centrale, gestandaardiseerde deel zich tot het decentrale deel?

Er zijn echter situaties denkbaar waarbij het, na voldoende inspanning en na het afleggen van het examen en de geoorloofde herexamens, een deelnemer niet lukt om de vereiste niveaus voor (alle) onderdelen van de talen en rekenen/wiskunde te behalen. In die gevallen moet er een mogelijkheid worden geschapen om, indien de deelnemer wél voldoende is beoordeeld voor de beroepsgerichte competenties, een certificaat af te geven om doorstroom naar de arbeidsmarkt mogelijk te maken. Op dit certificaat kunnen dan de behaalde niveaus voor de onderdelen van taal en rekenen/wiskunde worden vermeld. In dat geval kan een dergelijke deelnemer geen recht op doorstroom hebben.

Moet er ook een ondergrens worden gehanteerd voor deze vorm van certificering? Als een deelnemer wel slaagt voor zijn beroepsgerichte deel en daarmee terecht kan op de arbeidsmarkt (en dus ook in de maatschappij) moet er dan alsnog een ondergrens worden bewaakt voor de Nederlandse taalvaardigheid? En voor de Moderne Vreemde Talen? En voor rekenen/wiskunde? Een eventuele ondergrens voor Nederlands (bij wél een certificaat maar géén diploma) heeft wellicht niet zozeer een onderwijsinhoudelijke reden, maar meer een politiek-maatschappelijke: van inburgeraars wordt verwacht (geëist) dat ze het A2 niveau halen. Het ligt voor de hand te verwachten dat dit ook het absolute minimum voor Nederlands is voor alle mbo-kandidaten. Voor kandidaten op niveau 3 en 4 zou het denkbaar zijn de lat hoger te leggen, bijvoorbeeld voor Nederlands op B1 niveau.

In dit rapport spreken we ons niet uit over de compensatiemogelijkheden binnen taalvaardigheid en rekenen/wiskunde. Ook spreken we ons niet uit over het al dan niet hanteren van een absolute ondergrens en al zeker niet het aangeven van het niveau daarvan. We volstaan hier met erop te wijzen dat het hier gaat om een zeer gevoelige materie die nader onderzocht moet worden om hier een goed besluit over te kunnen nemen.

4 Gestandaardiseerde examens in een modern jasje

4.1 Inleiding

In hoofdstuk 3 hebben we modellen geschetst voor examinering van taal en rekenen/wiskunde. In deze modellen opteren we voor standaardisering van het examen door uit te gaan van een centraal ontwikkeld examen naast een decentraal, beroepsgericht deel waarvoor ook standaarden geformuleerd kunnen worden. In dit hoofdstuk gaan we nader in op de ontwikkeling van een centraal examendeel.

Centrale examinering roept bij sommigen een beeld op van conservatisme en starheid. Dat is onterecht; de toepassing van moderne testtheorieën biedt volop mogelijkheden om examensystemen te ontwikkelen die zowel betrouwbaar en valide (hét argument voor centrale examinering) als flexibel zijn. De moderne testtheorie heeft examenontwikkelaars de instrumenten gegeven om schalen te construeren waarmee cognitieve vaardigheden gemeten kunnen worden. Deze schaaltechnieken bieden verschillende voordelen ten opzichte van de meer traditionele toetstechnieken zoals bijvoorbeeld dat examenprestaties onderling beter vergelijkbaar zijn, ook als die gebaseerd zijn op examens die naar inhoud en niveau van elkaar verschillen. Zie bijlage 4 voor een beknopte toelichting op de inzet van schaaltechnieken voor onderwijskundige meetdoelen.

In dit hoofdstuk bespreken we de globale opzet van een gestandaardiseerd examensysteem dat gebaseerd is op schalingstechnieken en op de inzet van computers voor de afname¹⁴.

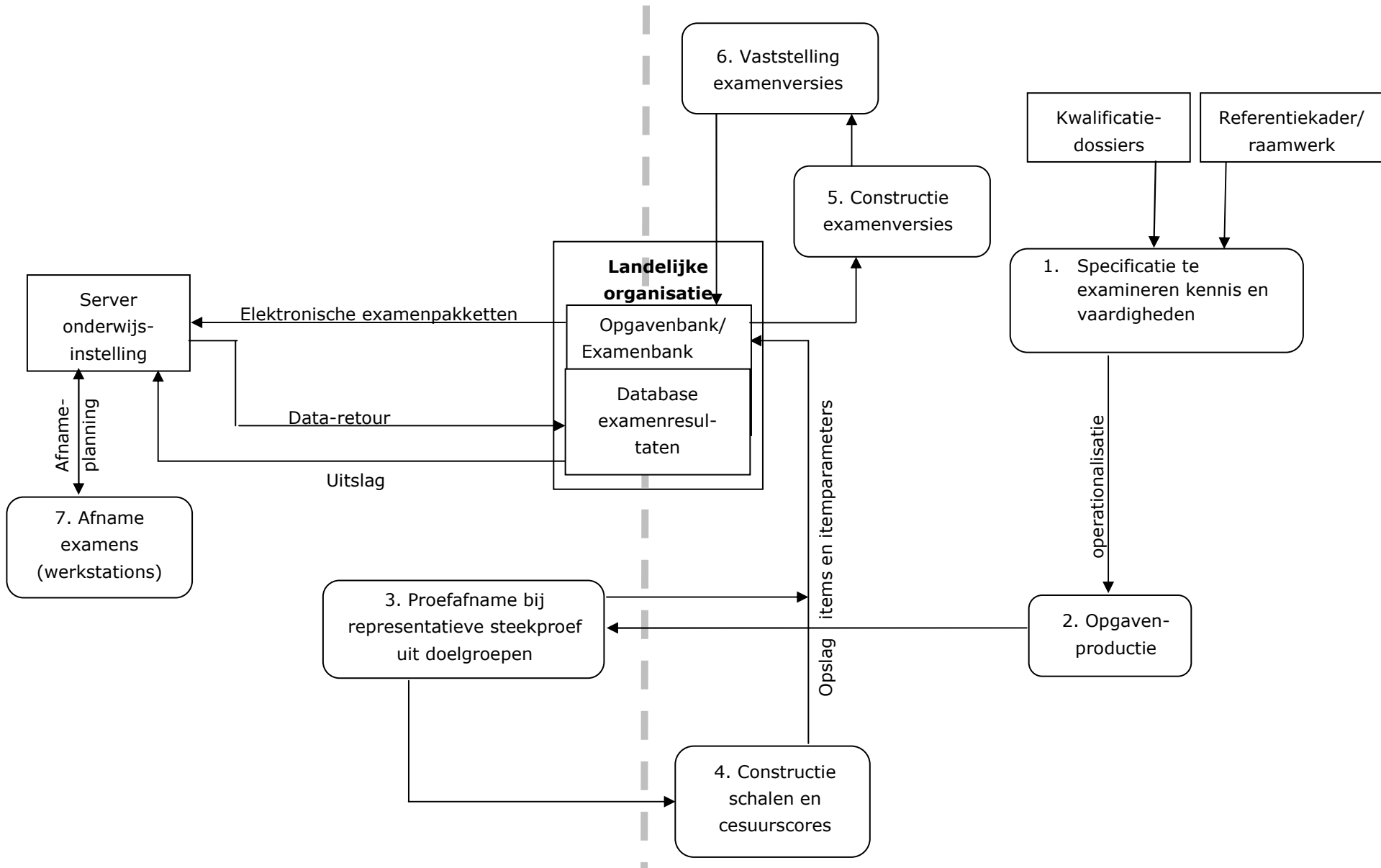
4.2 Opzet en werking

De bijgaande figuur (schets van de opzet en werking van een gestandaardiseerd examensysteem) geeft een overzicht van de belangrijkste activiteiten die uitgevoerd worden in de ontwikkel- en afname-omgeving van het examensysteem. In het hart van het schema zijn een opgavenbank/examenbank en een database met examenresultaten gepositioneerd. Deze centrale instrumenten worden beheerd door een onafhankelijke, landelijke organisatie onder regie van de overheid, die tevens een aansturende en coördinerende rol vervult voor de activiteiten die nodig zijn voor de ontwikkeling van de examens en de operationele afnames ervan. Hiervoor zijn verschillende varianten denkbaar zoals die nu ook in het onderwijs worden toegepast (bijvoorbeeld CEVO voor de examens in het voortgezet onderwijs of een staatsexamencommissie voor de Staatsexamens Nederlands als Tweede Taal). Vooralsnog volstaan wij hier met een meer algemene term.

¹⁴ De computer en internet kunnen ook worden gebruikt voor de ontwikkeling en distributie van de gestandaardiseerde formats voor het decentrale deel van het examen. Dit valt echter buiten het bestek van dit hoofdstuk.

Afname-omgeving

Ontwikkel-omgeving



4.2.1 *Specificatie van te examineren kennis en vaardigheden*

De kwalificatiedossiers van de verschillende uitstroomrichtingen per niveau en de referentiekaders of raamwerken van de vakgebieden Nederlands, een moderne vreemde taal en rekenen/wiskunde, vormen het uitgangspunt voor de specificatie van de kennis en vaardigheden die in aanmerking komen voor centrale examinering. Deze activiteit resulteert voor elk vakgebied of deelvaardigheid en daarbinnen per uitstroomrichting en per niveau in een specificatietabel met mogelijk te bevragen onderwerpen, al dan niet met een specificatie van de cognitieve verrichting (herinneren, begrijpen, toepassen) die wordt verwacht van de kandidaat ten aanzien van de genoemde onderwerpen.

4.2.2 *Opgavenproductie*

De specificatietabellen worden gebruikt door productieteams van docenten voor de constructie van examenopgaven. Het aantal opgaven dat per uitstroomrichting en per niveau geproduceerd moet worden is afhankelijk van het aantal examenmomenten¹⁵ dat men jaarlijks wil aanbieden én van het aantal opgaven per examen. Daarnaast moet rekening worden gehouden met uitval wegens niet functionerende opgaven (zie onder 4.2.3.) en met de noodzaak om reserve examenversies achter de hand te hebben. De opgaven worden gemaakt door ervaren docenten die affiniteit hebben met toetsconstructie. Docenten die toetreden tot een productieteam krijgen een korte cursus waarin ze de basisprincipes leren van het construeren van examenopgaven. De constructie vindt plaats onder leiding van professionele toetsontwikkelaars. Tot het moment dat de opgavenbanken een omvang hebben die garandeert dat de kans op verkeerde zak-/slaagbeslissingen door hergebruik van oude examenopgaven uiterst onwaarschijnlijk is, zal jaarlijks een portie nieuwe examenopgaven vervaardigd moeten worden die voldoende groot is om alle in het daarop volgende jaar in te zetten examenversies samen te stellen uit nieuwe examenopgaven.

4.2.3 *Proefafname*

Een implicatie van het gebruik van schalingstechnieken bij de centrale examinering van talen en rekenen/wiskunde in het mbo is dat examenopgaven op voorhand op kwaliteit worden onderzocht door ze te laten maken door een steekproef van kandidaten uit de doelgroep. Elke proefpersoon uit de steekproef krijgt een portie nieuw geconstrueerde opgaven te maken. Om te voorkomen dat kandidaten tijdens de proefafname opgaven moeten maken die ze ook tijdens het latere examen krijgen voorgelegd, loopt de proefafname minimaal één jaar voor op de operationele afname van het examen. Met de data die de proefafname oplevert, wordt de moeilijkheidsgraad van elke opgave geschat en wordt tevens een schatting gemaakt van de vaardigheid van de proefpersonen in het onderzoek. Met deze gegevens wordt vervolgens nagegaan of de geteste opgaven zich 'gedragen' volgens een bepaald (van tevoren gekozen) model. Alle opgaven die zich modelmatig gedragen worden geacht dezelfde vaardigheid te meten, wat ze in principe geschikt maakt voor gebruik in een examen. Alleen die opgaven worden opgenomen in de opgavenbank.

¹⁵ Vooralsnog wordt er vanuit gegaan dat elke examenversie volledig wordt samengesteld uit niet eerder gebruikte examenopgaven.

Het onderzoeksdesign dat wordt gebruikt, moet zodanig zijn dat de data die in een bepaald jaar verzameld worden, te verbinden zijn met data uit eerdere jaren. Op die manier kan een omvangrijke opgavenbank worden opgebouwd waarvan de opgaven allemaal op dezelfde schaal staan. Dat is een voorwaarde om op termijn over te kunnen gaan op computergestuurde (adaptieve) examinering.

4.2.4 Schaalconstructie en cesuurscores

Een schaal wordt verkregen wanneer de opgaven die voldoen aan het model worden geordend naar moeilijkheidsgraad. Op zo'n schaal kunnen de examenprestaties worden afgezet van kandidaten die een examen hebben gemaakt dat is samengesteld uit een zogeheten 'geschaalde opgavenbank'. Dat is alleen zinvol als er op diezelfde schaal ook een punt aanwijsbaar is dat minimaal bereikt moet zijn voor een positieve beslissing. Die punten heten cesuurscores. Het aanwijzen van een cesuurscore als de kwantitatieve samenvatting van wat iemand moet kennen en kunnen ten aanzien van een of ander domein is een lastige maar cruciale stap in de examenproductie. Voor de systematische ontwikkeling van dergelijke absolute prestatiestandaarden zijn in de literatuur verschillende procedures beschreven, waarvan de bekendste de Angoff-procedures¹⁶ zijn. Er zijn diverse varianten hiervan ontwikkeld waarvan de bekendste inhoudt dat elke deelnemer van een standaardbepalings-panel de waarschijnlijkheid op een correcte beantwoording aangeeft van elke opgave in het examen (of van elke opgave in een representatieve set opgaven uit de opgavenbank) door een kandidaat die met 'de hakken over de sloot' slaagt. De schattingen van de panel-leden worden gemiddeld, eventueel na een tweede ronde waarin getracht wordt grote geconstateerde verschillen tussen de panel-leden uit de eerste ronde te reduceren. De gemiddelde schattingen worden omgezet in een schaalwaarde die geldig is voor elk examen dat uit de betreffende geschaalde opgavenbank wordt samengesteld. Met deze of soortgelijke procedures kunnen op een bepaalde schaal de prestatiestandaarden voor verschillende onderwijsniveaus en/of doelgroepen of combinaties daarvan worden afgebeeld.¹⁷

4.2.5 Constructie examenversies

Examenconstructieteams zijn verantwoordelijk voor het vervaardigen van examenversies uit de nieuwe, nog niet eerder gebruikte opgaven. De teams letten met name op:

- *Inhoudelijke representativiteit.* De examenversies moeten inhoudelijk representatief zijn voor het vakgebied of deelvaardigheid daarbinnen.
- *Meetnauwkeurigheid.* De meetnauwkeurigheid van het examen moet het grootst zijn rondom de cesuurscore op de schaal. Met de beschikbare gegevens uit de proefafnames kan de meetkarakteristiek van elke examenversie vooraf in kaart worden gebracht én zo goed mogelijk worden afgestemd op het gebruiksdoel.
- *Niveau-handhaving.* Binnen een zelfde onderwijsniveau is het, vooral om psychologische redenen, gewenst dat de examenversies qua moeilijkheidsgraad niet te veel van elkaar verschillen. Vanuit psychometrisch oogpunt is dit niet per se nodig aangezien de scores op examens die zijn samengesteld uit dezelfde geschaalde opgavenbank een 'vertaling' krijgen naar een gemeenschappelijke

¹⁶ Cizek (2006), p. 239

¹⁷ Zowel Cito als Bureau ICE hebben de afgelopen jaren ervaringen opgedaan met het ontwikkelen van examens gekoppeld aan de CEF-niveaus.

vaardigheidsschaal en daardoor vergelijkbaar zijn, ook als ze naar inhoud en niveau aanzienlijk van elkaar verschillen.

4.2.6 Vaststelling examenversies

Een commissie die is samengesteld uit vertegenwoordigers van stakeholders controleert elke examenversie en stelt die, na een positief oordeel over de kwaliteit, vast.

4.2.7 Afname examens door instellingen

Het valt onder de verantwoordelijkheid van de instellingen om kandidaten x keer per jaar gelegenheid te geven om examen af te leggen. De tijdstippen worden voorgeschreven evenals de andere te nemen maatregelen die ervoor moeten zorgen dat de examens onder gestandaardiseerde condities worden afgenomen.

Enige tijd vóór een examenzitting kan een instelling een zogeheten 'examenpakket' downloaden vanaf de website van de landelijke examenorganisatie. Dit examenpakket bevat de examenversies van alle vakgebieden of deelvaardigheden daarbinnen voor de eerstvolgende examengelegenheid. Na installatie van het examenpakket worden de kandidaten, met behulp van de eveneens door de landelijke examenorganisatie beschikbaar gestelde software, ingepland voor één of meer examens (vakgebieden of deelvaardigheden daarbinnen).

Met een tijdens de examenzitting verstrekt wachtwoord krijgen de kandidaten toegang tot het (de) examen(s) waarvoor ze staan ingepland. Ook de software die gebruikt wordt om de examens op een beeldscherm af te nemen wordt verstrekt door de landelijke examenorganisatie. Kandidaten kunnen, indien gewenst en mogelijk, direct na afloop van het examen de (voorlopige) uitslag krijgen¹⁸.

De instellingen zijn verplicht om, met behulp van de functie 'dataretour' in de geleverde software, de afname-data ter beschikking te stellen van de landelijke examenorganisatie. Aldaar worden ze verwerkt en geanalyseerd met het oog op:

- verstrekking van officiële examenuitslagen aan de instellingen;
- rapportage over de landelijke examenresultaten;
- bijstelling itemparameters in de opgavenbanken.

4.3 Computergestuurde adaptieve toetsing (CAT)

In de vorige paragraaf is beschreven hoe met behulp van de computer geschaalde opgavenbanken gebruikt kunnen worden om daaruit conventionele examens samen te stellen van hoge psychometrische kwaliteit. Computer-based testing (CBT) heeft de laatste jaren een hoge vlucht genomen door enkele voor de hand liggende voordelen als vereenvoudigde logistiek, automatisch nakijken en scoren en de mogelijkheid om hogere cognitieve vaardigheden als probleemoplossen, beslissingen nemen, e.d. te beoordelen. Een interessante optie bij CBT is om de computer niet alleen te gebruiken om een toets of

¹⁸ Bij volledig gesloten examens die door de computer automatisch kunnen worden beoordeeld is dit een optie. Kiezen we voor een centrale examenvariant waarbij menselijke beoordeling van antwoorden nodig is, dan kunnen kandidaten niet direct de uitslag krijgen.

examen op een beeldscherm te presenteren en de antwoorden na te kijken en te scoren maar ook om die toetsen of examens door het computerprogramma te laten samenstellen uit een opgavenbank. Er zijn globaal drie manieren waarop dit kan:

1. Bij de eerste manier worden opgaven volgens een toevalsmechanisme uit de opgavenbank getrokken. Behalve op het aantal opgaven heeft men geen enkele controle op de toets die door het programma wordt samengesteld.
2. Bij de tweede manier probeert het programma zich te houden aan een toetsmatrijs die voorschrijft hoeveel opgaven over welke leerstofonderwerpen in de toets moeten worden opgenomen.
3. De derde manier, ten slotte, probeert die opgaven te kiezen die de meetfout van de toets zo klein mogelijk maken. Dit wordt bereikt door opgaven te kiezen die qua moeilijkheidsgraad zo goed mogelijk zijn afgestemd op de vaardigheid van de persoon die getoetst wordt. Deze derde manier wordt als uitgangspunt gehanteerd bij zogeheten computergestuurde adaptieve toetsen (CAT).

Een CAT werkt ongeveer als volgt :

Stap 1: Op het beeldscherm verschijnt een opgave.

Stap 2: De kandidaat geeft antwoord door een antwoord te kiezen uit de gepresenteerde alternatieven (bij meerkeuze vragen) of door zelf een (kort) antwoord te typen in het antwoordveld.

Stap 3: De software gaat na of het gegeven antwoord correct is.

Stap 4: De software maakt een schatting van de vaardigheid van de kandidaat op basis van alle tot dan correct en foutief beantwoorde opgaven.

Stap 5: De software bepaalt op grond van een 'stopregel' of de toets al beëindigd moet worden. Zo ja, dan volgt direct de rapportage over het behaalde resultaat. In het andere geval wordt uit de opgavenbank een opgave getrokken die qua moeilijkheidsgraad zo goed mogelijk is afgestemd op de geschatte vaardigheid van de kandidaat.

Na stap 5 gaat de procedure verder bij stap 2, totdat in stap 5 besloten wordt dat de toetsafname gestopt moet worden.

Naast de voordelen die van toepassing zijn op de inzet van geschaalde opgavenbanken (zie vorige paragraaf) biedt een CAT de volgende meerwaarde voor examineringsdoeleinden:

- Flexibilisering van het afnametijdstip. Omdat elke kandidaat in principe een andere toets maakt, is het risico op het aan elkaar doorgeven van exameninhouden vrij beperkt. Hierdoor wordt het mogelijk om veel vaker dan nu gebruikelijk is een examen af te leggen.
- Altijd het juiste niveau. Een adaptief examen is nooit te moeilijk of te gemakkelijk voor een kandidaat. Dat is met name in het MBO een groot voordeel aangezien de deelnemers vaak een grote heterogeniteit vertonen met betrekking tot de te meten kennis en vaardigheden. Een adaptief examen wordt zodanig samengesteld dat met de grootst mogelijke meetnauwkeurigheid bepaald wordt waar op de schaal de vaardigheid van de kandidaat ligt.

- Onmiddellijke uitslag. Uit de hierboven beschreven procedure is duidelijk geworden dat de CAT-software tijdens het examen de antwoorden van de kandidaat nakijkt en scoort. Daardoor kan onmiddellijk na de beantwoording van de laatste opgave de (voorlopige) uitslag aan de kandidaat meegedeeld worden. De officiële uitslag krijgen instellingen zodra de data retour zijn gezonden naar de organisatie die zorg draagt voor opslag en verwerking van de examenresultaten.
- Organisatorische eenvoud. Op elk door de school gewenst moment zijn de examens klaar om afgenomen te worden. Het enige waarvoor gezorgd moet worden is toezicht.

CAT is inmiddels zo ver ontwikkeld dat het kan worden ingezet voor het nemen van zwaarwegende beslissingen over kandidaten. Een tastbaar bewijs daarvoor is de instroomtoets 'rekenen' die vanaf studiejaar 2006-07 verplicht is voor alle eerstejaars pabo-studenten. Elke student krijgt gedurende het eerste studiejaar drie kansen. Op grond van de derde kans wordt een bindend studie-advies gegeven. De ervaringen van docenten, studenten en andere belanghebbenden gedurende het eerste operationale jaar waren over het algemeen zeer gunstig.

CAT kan alleen als gekozen wordt voor een centraal examen dat volledig bestaat uit gesloten vragen.

Ondanks deze aansprekende voordelen zijn wij van mening dat CAT pas op de langere termijn een serieuze optie kan zijn voor de centrale examinering van MBO-kandidaten met betrekking tot de hierboven aangegeven vakgebieden. De reden hiervoor is dat grote aantallen kandidaten die niet gelijktijdig maar na elkaar geëxamineerd worden zeer omvangrijke opgavenbanken vergen om geheimhouding van examenmateriaal in voldoende mate te kunnen garanderen. Men moet daarbij denken aan enkele duizenden opgaven per te examineren vakgebied of deelvaardigheid daarbinnen. Dit is alleen haalbaar wanneer de ontwikkeling van dergelijke banken over langere tijd kan worden gespreid. Inhoudelijk lijkt dat weinig bezwaarlijk aangezien het bij genoemde vakgebieden gaat om stabiele leerstof die weinig last heeft van veroudering. Belangrijk in dit verband is dat bij de keuze en uitwerking van een tussenoplossing wel al rekening moet worden gehouden met een toekomstige overgang naar computergestuurde adaptieve examinering.

5 Aanbevelingen behorende bij dit advies

5.1 Bepaal de niveaus voor de talen opnieuw

Bij de eerste versie van de taalcompetentieprofielen zijn twee aannames gedaan:

1. Er is voor Nederlands gebruik gemaakt van het Raamwerk NT2 (voor tweede taalleerders) bij gebrek aan een specifiek raamwerk Nederlands. Er is aanbevolen om deze niveaus te herzien wanneer er een specifiek Raamwerk Nederlands beschikbaar is.
2. In de eerste versie is nog uitgegaan van het principe van 'schering en inslag' en zijn de taalniveaus ook vastgesteld met het oog op de doorstroom. Als gevolg hiervan zijn in sommige gevallen erg hoge niveaus aangegeven (bijvoorbeeld C1).

Nu er een specifiek Raamwerk Nederlands is en de minimale niveaus Nederlands voor Leren, Loopbaan en Burgerschap voor alle deelnemers zijn vastgesteld is het van groot belang dat de Kenniscentra opnieuw de niveaus voor Nederlands vanuit de eisen van het beroep vaststellen. Hiervoor moeten zij de taalniveaus werkelijk in de praktijk gaan onderzoeken en daarbij gebruik maken van het nieuwe Raamwerk Nederlands.

Deze zelfde exercitie zal ook moeten gebeuren voor de Moderne Vreemde Talen, daar waar geen onderzoek in de praktijk is gedaan.¹⁹

5.2 Zorg voor goede taalassessoren

Taalvaardigheid beoordelen is een specialisme. De 'gewone' certificering tot assessor is niet geschikt voor taalassessoren. Bovendien kunnen taaldocenten geen gecertificeerd assessor worden omdat ze niet gekoppeld zijn aan één kwalificatie maar vaak in veel verschillende opleidingen lesgeven en dus ook moeten examineren.

We adviseren om specifieke eisen te stellen aan taalassessoren. Speciale trainingen kunnen worden aanbevolen.

5.3 Taal mag geen struikelblok zijn

Naast het expliciet versterken van het Nederlands in het mbo is het ook nodig dat zowel het lesmateriaal als de examenopgaven afgestemd zijn op het taalniveau van de deelnemer. Te vaak is dit in te moeilijk Nederlands gesteld en levert dit een oneigenlijk struikelblok op bij het leren. Bovendien werkt lesmateriaal dat te ver van het beheersingsniveau van de deelnemer afligt, niet bevorderend voor het verwerven van taal. Om taalverwerving tijdens de beroepsopleidingen te vergroten, is het van belang dat lesmateriaal en taalniveau met elkaar matchen.

¹⁹ Sommige Kenniscentra hebben dit wél zeer gedegen onderzocht en vastgesteld.

Wij adviseren om een specifieke kwaliteitsstandaard op te nemen die moet garanderen dat het taalniveau van een examen passend is bij het (verwachte) taalniveau van de deelnemer.

5.4 Stel minimale niveaus voor Rekenen/Wiskunde voor alle deelnemers in het kader van Burgerschap verplicht

Dit is ook een van de aanbevelingen uit het Raamwerk Rekenen-Wiskunde. Dit zijn dus géén niveaus voor de doorstroom.

5.5 Bepaal minimale niveaus voor Rekenen/Wiskunde per doorstroomvariant

Omdat de doorstroomeisen voor Rekenen/Wiskunde zo divers zijn moet dit per doorstroomvariant worden uitgewerkt. Hiervoor is het nodig om binnen de beroepspraktijk én de vervolgopleidingen onderzoek te doen naar de aard van rekenen/wiskunde taken.

5.6 Bepaal de vereiste niveaus voor Rekenen/Wiskunde per kwalificatieprofiel

Voor rekenen/wiskunde moet onderzoek worden gedaan naar het karakter van "workrelated mathematics" en hoe dat zich laat vertalen naar onderwijs. Per kwalificatieprofiel moet een rekenen/wiskunde profiel worden opgesteld (zie ook Bijlage 3).

5.7 Stel een beslismodel op

In hoofdstuk 3 wezen we al op de afwegingen die moeten worden gemaakt bij het al dan niet toekennen van een voldoende voor taal en rekenen/wiskunde. Het nadenken over compensatiemogelijkheden tussen verschillende onderdelen binnen taal en binnen rekenen/wiskunde is een voorwaarde om het voorgestelde examenmodel te kunnen implementeren. Ook het al dan niet invoeren van een ondergrens en het bepalen van waar de ondergrens zou moeten komen te liggen, is een discussie die beslecht moet worden alvorens tot de invoering van voorgestelde examenmodellen over te gaan.

5.8 Betrek dit advies bij de ontwikkelingen van examenprofielen

COLO ontwikkelt op dit moment voorbeelden van examenprofielen bij kwalificatieprofielen, in opdracht van de staatsecretaris. Dit traject loopt al. Zorg er voor dat de adviezen uit dit rapport hiermee worden afgestemd en omgekeerd.

Literatuur

- Bohnen, E. (et al) (2007) *Raamwerk Nederlands. Nederlands in (v)mbo-opleiding, beroep en maatschappij*. De Bilt: MBO-raad
- Cizek, G.J. (2006). Standard setting. In: S.M. Downing & T.M. Haladyna (Eds.), *Handbook of Test Development*. Mahwah (NJ): Lawrence Erlbaum Ass, Publishers.
- Council of Europe (2001) *Common European Framework of Reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press
- Driessen, M. en Edelenbos, P. (2006) *Talen tellen: taalcompetentieprofielen in de nieuwe kwalificatiestructuur*. 's-Hertogenbosch: CINOP
- Driessen, M. (red.) (2007) *Referentiedocument 'Talen in de kwalificatieprofielen'. Moderne vreemde talen en Nederlands. Herziene versie 2.0*. De Bilt: CINOP/MBO-raad
- Driessen, M. , A. van Kleef en A. Fitzpatrick (2007) *Talen toetsen: van taalcompetentieprofielen naar bewijzen van bekwaamheid*. 's-Hertogenbosch: CINOP
- Gemeenschappelijk Europees Referentiekader voor moderne vreemde talen: Leren, Onderwijzen, Beoordelen* (2006), Den Haag: Nederlandse Taalunie
- Hoyle, C.&Noss, R. (1998). Anchoring Mathematical Meanings in Practice. In: K. Gravemeijer, ed. *Introductory texts for the International Conference on Symbolizing and Modeling in Mathematics education*. Freudenthal Instituut, Utrecht.
- Inspectie van het Onderwijs (2007). *Toezichtskader examinering 2007-2008*. Concept versie. Utrecht: Inspectie van het Onderwijs.
- LICA (september 2006) *Onderzoeksverslag Duurzaamheid en versterking doorstroom mbo-hbo 'Pontons tussen twee bewegende oevers'*
- Liemberg, E. (2001) *Raamwerk Moderne Vreemde Talen*. De Bilt: Bve-raad
- Liemberg, E., Driessen, M. en van Kleunen, E. (2004). *Evaluatie Taalportfolio BVE 2001-2004*. 's-Hertogenbosch: CINOP
- MBO Raad (2007) *Standpunt.mbo: Elke deelnemer leert Nederlands op vereist niveau*. (op www.mboraad.nl, 8 jan. 2008)
- Neuvel, J., Bersee, T., Exter, H. den, & Tijssen, M. (2004). *Nederlands in het middelbaar beroepsonderwijs. Een verkennend onderzoek naar het onderwijsaanbod Nederlands en de taalvaardigheid van leerlingen*. 's-Hertogenbosch: Cinop.
- Noijons, J. en Kuijper, H. (2006) *Koppeling van de centrale examens leesvaardigheid moderne vreemde talen aan het Europees Referentiekader*. Arnhem: Cito

Onderwijsraad (2006) *Versteving van kennis in het onderwijs*. Den Haag: Onderwijsraad

Steen, L.A. (2003), *Data, Shapes, Symbols: Achieving Balance in School Mathematics. Quantitative Literacy: Why Numeracy Matters for Schools and Colleges*, Bernard Madison and Lynn Arthur Steen, editors. National Council on Education and the Disciplines, 2003, pp. 53-74.

Straetmans, G.J.J.M., & Eggen, T.J.H.M. (1998). Computerized Adaptive Testing: What it is and how it works. *Educational Technology*, 38, 1, 45-52.

Straetmans, G.J.J.M., & Eggen, T.J.H.M. (2005). Afrekenen op rekenen: over de rekenvaardigheid van pabo-studenten en de toetsing daarvan. *Tijdschrift voor Hoger Onderwijs*, 23, 3, 123-139.

Straetmans, G.J.J.M., & Eggen, T.J.H.M. (2007). WISCAT-pabo: Een computergestuurd adaptief toetspakket rekenen. *Onderwijsinnovatie*, 3, 17-27.
http://www.ou.nl/Docs/TijdschriftOI/OI_3_07.pdf

TWIN (1998). Rapport van de werkgroep Doorstroming MTO-HTO voor de vakken wiskunde en natuurkunde. Freudenthal Instituut.

Vispoel, W.P., Rocklin, T.R., & Wang, T. (1994). Individual differences and test administration procedures: a comparison of fixed-item, computerized-adaptive, and self-adapted testing. *Applied Measurement in Education*, 7, 53-79.

Weijers et al. (2007) *Raamwerk rekenen-wiskunde voor het mbo*. Utrecht: Freudenthal Instituut.

Bijlage 1 Onderzoeken naar taal- en rekenvaardigheden

Veel mbo-gediplomeerden maken gebruik van het feit dat hun diploma op niveau 4 toegang verleent tot een studie in het hbo. Inmiddels is zo'n dertig procent van alle instromers afkomstig uit het mbo. Meer dan de helft daarvan rondt de hbo-studie na vijf jaar af met een diploma. De rest loopt ernstige vertraging op, switcht naar een andere studie of stopt er helemaal mee. Volgens HBO-instellingen wordt dit veroorzaakt door tekorten op het gebied van abstract denken en taal- en rekenvaardigheden. Op het gebied van rekenen kon het veronderstelde tekort aan kennis en vaardigheid bevestigd worden door resultaten uit onderzoek. In een grootschalig onderzoek onder eerstejaars pabo-studenten²⁰ bleek dat bijna driekwart van de studenten die afkomstig waren van het mbo, niet voldeed aan de instroomeis volgens welke een instromende pabo-student net zo goed moet kunnen rekenen als een goede leerling uit groep 8. Dit probleem wordt al een aantal jaren gesignaleerd vanuit de Pabo's zelf. Het wordt voor een groot deel veroorzaakt door de leerlingen die geen reken/wiskundeonderwijs meer hebben gehad na klas 2 van het vmbo (voornamelijk in de sectoren zorg en welzijn en economie). Het betekent dat een percentage van 20% van de vmbo-uitstroom na hun veertiende geen reken/wiskundeonderwijs meer heeft genoten. Het is met name deze categorie leerlingen die via de 'sluiproute' van het mbo wil binnenkomen op de Pabo. In september 2006 verscheen het Lica rapport over de doorstroom van MBO naar HBO. Hierin wordt melding gemaakt van verschillende soorten problemen bij die doorstroom. Een van de problemen die als kenmerkend worden beschreven is het gebrek aan basiskennis en basisvaardigheden bij de MBO student die doorstroomt naar het HBO. Hierbij worden expliciet genoemd: Nederlandse taalvaardigheid (en onderdelen daarvan zoals stellen, spellen, samenvatten), de moderne vreemde talen en elementaire rekenvaardigheid (en ook het interpreteren van grafieken en tabellen, het maken van eenvoudige vergelijkingen). Verder constateert het Lica rapport dat er in de nieuwe kwalificatieprofielen voor het MBO 'niet of nauwelijks inhoudelijk en/of structurele aanzetten worden gegeven tot competenties die zich richten op de doorstroom van mbo-4-deelnemers naar het hbo en dat er dus ook geen werkbare suggesties worden gegeven voor relevante leercompetenties die voor doorstromers van groot belang zijn. Gelet op de meer dan 60% doorstromers van niveau 4 naar het hbo is dit – in de huidige situatie met alle ervaringen die de afgelopen tien jaar zijn opgedaan - naar het oordeel van de onderzoekers ronduit verbijsterend' (pag. 24).

Het opzetten van leerlijnen voor deze basisvaardigheden wordt 'geen overbodige luxe', maar een 'bittere noodzaak' genoemd. In het kwalitatieve onderzoek (IVA Beleidsonderzoek en Advies (augustus 2006) 'Oplossingen voor kennisdeficiënties', in opdracht van de Onderwijsraad, naar kennisdeficiënties wordt in een aantal casussen waarbij het mbo (en ook het toeleverende vmbo en het afnemende hbo) is betrokken vastgesteld dat veel leerlingen tekorten hebben in onder andere rekenvaardigheid en taalvaardigheid. De onvrede hierover kwam niet alleen van de scholen zelf maar ook van het bedrijfsleven. Daarnaast vormen deficiënties in de

²⁰ Straetmans & Eggen, 2005.

zogenaamde 'sociale' sleutelvaardigheden zoals het vermogen tot zelfstandig werken, tot samenwerken, op tijd komen, voor alle leerlingen bij de aansluiting vmbo-mbo het grootste probleem. Bij de aansluiting van het mbo met het technische hoger onderwijs ontbreekt het studenten aan reken- en wiskundige vaardigheden. Vooral het gebrek aan het vermogen tot abstraheren breekt de studenten op. In het recent verschenen Raamwerk rekenen/wiskunde mbo²¹ wordt daarover het volgende opgemerkt:

Bij de doorstroom vanuit het mbo (niveau 4) naar het hbo in een aanverwante opleiding (bijvoorbeeld van mbo-werktuigbouw naar hbo-werktuigbouw) doet zich het probleem voor dat rekenen/wiskunde die nodig is om het (mbo-)beroep uit te oefenen niet altijd spoort met rekenen/wiskunde die nodig is voor de instroom in de hbo-opleiding. Om deze reden was het in het mbo een goede praktijk om voor doorstromers naar het hbo extra rekenen/wiskunde aan te bieden, naast het op het mbo-diploma gerichte aanbod in rekenen/wiskunde. Tot 2002 gebeurde dit middels de verplichte doorstroomprogramma's wiskunde en natuurkunde voor de sector techniek, die werden afgerond met een landelijke eindtoets. Deze verplichting kwam te vervallen met het politieke besluit om elke mbo-deelnemer doorstroomrecht te verlenen naar het hele hbo, ook voor niet-verwante opleidingen. Voor de doorlopende leerlijn maakt het niet uit waar die extra wiskunde gegeven wordt, in het mbo of in het hbo, als er maar rekening mee wordt gehouden dat er voor een aantal hbo-opleidingen meer wiskunde nodig is dan voor de beroepskwalificatie op mbo niveau 4 vereist is. Per opleiding of per cluster opleidingen zouden hierover afspraken gemaakt moeten worden. Het raamwerk rekenen/wiskunde kan daarbij een bruikbaar instrument zijn. De discrepanties in het beheersingsniveau kunnen ermee zichtbaar en bespreekbaar gemaakt worden.

Bij doorstroom naar andere hbo opleidingen komt ook de gebrekkige taalvaardigheid weer aan de orde. Ook Engels wordt hierbij genoemd. Deze opvatting wordt ondersteund door onderzoek dat in het mbo gedaan werd naar het beheersingsniveau van Nederlands en waaruit geconcludeerd werd dat meer dan de helft van de deelnemers over onvoldoende taalvaardigheid in het Nederlands beschikt om op school en in de beroepspraktijk adequaat te functioneren.

In 2004 voerde CINOP een onderzoek uit naar de Nederlandse taalvaardigheid van MBO-deelnemers. Uit dit onderzoek kwam naar voren dat het taalniveau van meer dan de helft van de leerlingen een half tot een heel niveau lager ligt dan het taalniveau dat nodig is om op school en in de beroepspraktijk naar behoren te functioneren. Hoewel het onderwerp rekenen/wiskunde geen onderdeel uitmaakte van dit onderzoek, werd bij de interviews met ROC-docenten spontaan veelal ingebracht dat hetzelfde ook geldt voor reken/wiskunde-vaardigheden. Dat was de aanleiding om in samenspraak tussen CINOP en FI de reken/wiskunde-vaardigheden onder de aandacht van het BVE-veld te brengen, hetgeen uiteindelijk is uitgemond in het formuleren van een Raamwerk rekenen/wiskunde mbo (december 2007).

Grootschalig onderzoek dat Bureau ICE in de zomer van 2007 uitvoerde²², bevestigde het beeld. Daaruit bleek dat voor leesvaardigheid een kwart van de

²¹ Raamwerk rekenen/wiskunde mbo, M. Wijers e.a., Freudenthal Instituut, versie 0.9, december 2007

instromende MBO-ers niet verder komt dan taalniveau A2.²³ Ruim de helft van de instromende MBO-ers komt op het B-niveau Lezen binnen en nog geen tien procent op taalniveau C1.

Dit is zorgwekkend gezien de ambities die geformuleerd zijn in het document Leren, Loopbaan en Burgerschap. En zorgwekkend omdat zowel in het competentiegerichte onderwijs als in de maatschappij vaker en meer een goede communicatieve taalvaardigheid vereist is.

Tegelijkertijd wees het onderzoek van Cinop uit dat het onderwijsaanbod om Nederlands te leren tekort schiet. Dit geldt zowel voor de aparte lessen Nederlands als voor het (deels) geïntegreerde aanbod. Ook het rapport van de Inspectie voor het Onderwijs liet zien dat er inderdaad maar beperkt tijd wordt besteed aan het vak Nederlands in het beroepsonderwijs. Hoewel er geen exacte urentabellen beschikbaar zijn voor het maken van een niveau-stap lijkt de hoeveelheid tijd die nu beschikbaar is voor taal gekoppeld aan de kloof die er is tussen het taalniveau dat deelnemers hebben en zouden moeten bereiken, dat zorg op zijn plaats is.

Voor rekenen/wiskunde zijn tot nu toe geen diepgravende studies uitgevoerd om de (on)gecijferdheid in de beroepskolom echt in kaart te brengen. Een van de redenen daarvoor is dat de politiek tot voor kort (begin 2006) daar geen interesse voor toonde. De publieke discussie rond de vermeende tekorten op dit gebied voor beginnende Pabo-studenten (begin 2006) heeft echter veel in gang gezet.

Gerichte onderzoeken in Nederland naar mogelijke oorzaken zijn echter nog niet uitgevoerd. De kreet “ze kunnen niet meer rekenen” is te algemeen, en daarmee nietszeggend, om alleen daarop de alarmbel te luiden. Verder onderzoek is zeker nodig waarbij deze algemene kreet meer wordt gespecificeerd in de richting van beroepsspecifieke- en burgerschaps vaardigheden. De enige echte onderzoeksgegevens van dit moment komen uit internationale onderzoeken zoals PISA en TIMSS.

Daarbij blijkt Nederland het goed te doen, zowel op het gebied rekenen/wiskunde, als ook bij science en taal. Voor rekenen/wiskunde behoort Nederland tot de absolute top: bij PISA 2003 bezette Nederland de vierde plaats (na Hong Kong, Finland en Zuid-Korea op een totaal van 40), evenals in 2006 de vijfde plaats (na Taipei, Finland, Hong Kong en Zuid-Korea op een totaal van 57 landen) in een landen-totaal van rond de 40. De politiek roept dan ook onmiddellijk dat er dus geen probleem is Tegelijkertijd zijn er op nationaal niveau allerlei gremia (waaronder diezelfde politiek) die roepen dat het slecht gesteld is met het niveau van rekenen/wiskunde. Eén van de mogelijke oorzaken daarvan kan worden gezocht in de manier waarop de resultaten van internationaal onderzoek worden gerelateerd (of juist niet) aan nationale standaarden. De PISA-scores zijn gerelateerd aan de mediaan-score van alle deelnemende landen. Daarmee wordt een relatieve score (de mediaan wordt gesteld op de score 500) geïntroduceerd. Nederland scoort dan “ruim” boven deze mediaan-score (531). De nationale standaard lijkt echter te zijn gebaseerd op het idee dat Nederland een kenniseconomie moet zijn, waarbij volgens de Lissabon-accorden 50% van de populatie tenminste het hbo-niveau dient te bereiken. Een land als Brazilië heeft deze doelstelling niet. Daarmee wordt

²² Rapport verschijnt in januari 2008

²³ Zie voor een beschrijving van de taalniveaus hoofdstuk *

de lat voor het Nederlandse onderwijs op het gebied van rekenen/wiskunde veel hoger gelegd dan in veel andere landen.

Naar aanleiding van de resultaten van PISA-2003 is door het Freudenthal Instituut en het Cito een analyse uitgevoerd waarbij deze nationale criteria zijn gelegd naast de prestaties van BB- en KB-leerlingen van het VMBO. Daarbij blijkt onder andere dat deze categorie leerlingen wel goed is op het onderste niveau (Reproductie), maar slecht scoort op de hogere niveau's (Connections en Reflection) die in het kader van de nationale doelstellingen voor het onderwijs belangrijker zijn. Kennelijk is de procedurele (op algoritmen gebaseerde) kennis wel toereikend bij deze categorie leerlingen, maar het *functioneel hanteren* van deze kennis in contextueel diverse toepassingsgebieden blijft daar ver bij achter. En deze vaardigheden vormen kennelijk binnen welke beroepscontext dan ook de belangrijkste criteria om te constateren dat de vaardigheden onder de maat zijn. Het lijkt er op dat kennis op zichzelf niet een garantie biedt voor de hogere vaardigheid van het functioneel hanteren daarvan binnen een (nieuwe) context. Daarom luiden een paar van de voorzichtig geformuleerde conclusies van het Nederlandse PISA-onderzoek²⁴:

Er dient meer uitdagend en authentiek probleem georiënteerd wiskundeonderwijs te komen op het vmbo. Dit omdat het niet verantwoord is binnen de Nederlandse context dat zoveel vmbo-leerlingen het geletterdheidsniveau 4 niet halen. Daarbij dient met name ook het taalaspect van deze wiskunde aan de orde te komen: het lezen en interpreteren van een tekst en visuele informatie en het verwoorden en in beeld brengen van de oplossing.

Er dient nader onderzoek te komen naar het feit dat het niveau van het Nederlandse wiskundeonderwijs lijkt te dalen als we de gangbare internationale onderzoeken PISA en TIMSS als maatlat hanteren. Daarbij lijkt het verstandig het niet te laten bij een verwijzing naar de teloorgang van de staartdeling.

En:

Dit vereist deels een ander type opdrachten dan die in de meeste schoolboeken staan. In de Nederlandse wiskundeboeken komen veel opdrachten voor die bestaan uit deelvragen waarop een kort antwoord voldoet. Er komt zelden een echte complexe probleemstelling voor en er worden weinig redeneringen gevraagd. Er wordt niet echt gewerkt aan de wiskundige geletterdheid zoals die door PISA is gedefinieerd, en zoals die voor het halen van het PISA-niveau 4 gewenst is.

Uit het feit dat de Nederlandse leerlingen in het algemeen, en vmbo-leerlingen in het bijzonder, slecht scoren op de vragen uit het cluster Reflectie en op het gehele domein probleemoplossen, concluderen we dat er behoefte is aan authentieke probleemstellingen waarin dit type hogere orde vaardigheden toegepast moet worden.

Er liggen kansen voor betekenisvolle opdrachten in de directe schoolomgeving van de leerlingen door bijvoorbeeld aan te sluiten bij de sectorspecifieke beroepsgerichte vakken en daarin de wiskunde te gebruiken om een probleem op te lossen.

²⁴ Wiskundige geletterdheid volgens PISA, hoe staat de vlag erbij? Deel 1. Analyse. Jan de Lange e.a., Freudenthal Instituut, 2006

Er zou minder nadruk kunnen worden gelegd op het geven van een antwoord alleen. Leerlingen zouden moeten leren dat een eerste poging niet altijd meteen goed hoeft te zijn, dat het maken van schetsjes en het gebruik van kladpapier iets is wat helpt (en mag).

Er zou voor de vmbo-leerlingen meer aandacht gegeven moeten worden aan de taal, zowel aan het lezen, spreken en begrijpen als aan het schrijven, bijvoorbeeld bij het verwoorden en opschrijven van het oplossingsproces. Daarbij zouden meer verbindingen tussen alledaagse taal en wiskundige begrippen moeten worden gelegd.

Bij al het bovenstaande geldt dat er meer gelegenheid zou moeten zijn voor onderlinge uitwisseling (interactie) en reflectie. Er zou wat minder nadruk kunnen liggen op het individueel maken en nakijken van sommen, want dat lijkt toch wat anders te zijn dan wiskundige geletterdheid. Elkaars oplossingen en verwoordingen bekijken, bespreken en beoordelen kan een goede aanvulling in die richting zijn.

In het vmbo geldt nog dat wiskunde een zelfstandig vak is. De aanbeveling om meer authentieke beroepscontexten te gebruiken bij het leren van rekenen/wiskunde geldt natuurlijk in het competentiegerichte MBO veel minder. In het MBO ligt het probleem eerder tegenovergesteld: het leren vindt plaats binnen de beroepscontext met zijn beroepstaken. Het blijkt beslist niet gemakkelijk om daarbinnen specifieke reken- en wiskundige kennis en vaardigheden aan te wijzen die in het kader van die beroepscontext zijn aan te leren. Vaak gaat het om het gebruiken van tamelijk elementaire kennis en vaardigheden die zijn gemengd met de beroepsspecifieke kenmerken van het probleem. In specifieke gevallen gaat het ook om een uitbreiding van wiskundige kennis binnen het betreffende vakgebied.

Voor de beroepskwalificatie lijkt het zaak om al bekende (procedurele) kennis op het gebied van rekenen/wiskunde voor de deelnemer herkenbaar te koppelen aan de specifieke probleemstellingen in de beroepscontext. Daarbij zal de aanbieder goed op de hoogte moeten zijn van de methoden en technieken zoals die tegenwoordig worden gehanteerd in het vmbo (zoals de verhoudingstabel, verschillende methoden voor het oplossen van vergelijkingen, het werken met verhoudingen en procenten).

Bijlage 2 Threshold Level

Een belangrijk uitgangspunt bij de bepaling van de minimale uitstroomniveaus voor Nederlands is geweest het niveau van een onafhankelijke taalgebruiker: het zogenaamde 'Threshold Level', dat in het ERK het niveau B1 is.

Niveau B1 is in het oorspronkelijke ERK als volgt globaal omschreven:

'Kan de belangrijkste punten begrijpen uit duidelijke standaardteksten over vertrouwde zaken die regelmatig voorkomen op het werk, op school en in de vrije tijd. Kan zich redden in de meeste situaties die kunnen optreden tijdens het reizen in gebieden waar de betreffende taal wordt gesproken. Kan een eenvoudige lopende tekst produceren over onderwerpen die vertrouwd of die van persoonlijk belang zijn. Kan een beschrijving geven van ervaringen en gebeurtenissen, dromen, verwachtingen en ambities en kan kort redenen en verklaringen geven voor meningen en plannen.'

Uit het Raamwerk Nederlands:

Luisteren

Begrijpt concreet geformuleerde uitleg en informatie over veel voorkomende onderwerpen uit opleiding en werk.

Lezen

Kan concreet geformuleerde teksten van de opleiding, het werk of uit het dagelijks leven lezen met een redelijke mate van begrip. De teksten hebben een duidelijke opbouw met veelal frequent woordgebruik.

Gesprekken voeren

Kan gesprekken voeren over alledaagse en niet alledaagse onderwerpen uit het dagelijkse leven, werk en opleiding.

Kan uiting geven aan persoonlijke meningen, kan informatie uitwisselen over vertrouwde onderwerpen.

Spreken

Kan redelijk vloeiend ervaringen, gebeurtenissen, meningen, dromen, verwachtingen, gevoelens en ambities redelijk helder beschrijven.

Schrijven

Kan heldere samenhangende teksten schrijven met een simpele, lineaire opbouw, over uiteenlopende vertrouwde onderwerpen binnen zijn werk en opleiding.

Het B1 niveau kan hiermee gezien worden als een niveau passend bij een startkwalificatie, een MBO niveau 2 opleiding, niet alleen of specifiek voor het beroep maar vooral ook voor de kansen op doorstroom naar een vervolgopleidingen en het functioneren in de maatschappij.

Bij de ontwikkeling en de beschrijving van de niveaus van het ERK is gekeken naar de kwaliteit en de kwantiteit van taalvaardigheid. Met kwaliteit wordt uitgedrukt op

welke dimensies de taalvaardigheid zich kan ontwikkelen, zoals uitspraak, coherentie, grammaticale correctheid en de omvang van de woordenschat. Met kwantiteit wordt bedoeld het aantal domeinen, functies, noties, rollen etc. waarmee de taalgebruiker uit de voeten kan. Taalvaardigheid ontwikkelt zich dus altijd multi-dimensionaal en niet noodzakelijkerwijs lineair. De ontwikkeling van taalvaardigheid (vooral die van de moedertaal) heeft een sterke relatie met de algemene ontwikkeling en persoonlijkheidskenmerken zoals intelligentie en motivatie. In het ERK wordt dit expliciet verwoord in hoofdstuk 5 (De competenties van de taalgebruiker/taalleerder). Hierin wordt beschreven hoe in feite elke menselijke competentie op de een of andere manier bijdraagt aan de vaardigheid tot communiceren en omgekeerd.

De algemene competenties van een taalgebruiker die nauw verweven zijn met communicatieve competenties (en dus met taalcompetenties) hebben te maken met, onder andere kennis van de wereld, sociaal-culturele kennis, existentiële competentie (persoonlijkheidsfactoren) en leervermogen, waaronder studievaardigheden.

Bij de ontwikkeling van het ERK voor het leren van een tweede of vreemde taal is men ervan uitgegaan dat deze kennis en vaardigheden bij een volwassene min of meer aanwezig zijn. Dit is natuurlijk bij kinderen en jongeren (nog) niet het geval. Daar verloopt de ontwikkeling van deze algemene competenties min of meer gelijk op met de ontwikkeling van communicatieve taalcompetenties.

Ook mensen (volwassenen of jongeren) die achterstanden of deficiënties hebben in deze kennis en vaardigheden zijn ook in hun communicatieve competenties minder ontwikkeld. Het verder ontwikkelen van hun taalcompetenties zou dan gepaard moeten gaan met een intensieve aandacht voor de verdere ontwikkeling van hun denkvaardigheden.

In het nieuwe Raamwerk Nederlands voor het (v)mbo is om deze reden in het format van de niveau-beschrijvingen een extra kolom 'kennis' opgenomen. Deze is onderverdeeld in:

- Kennis van het onderwerp
- Kennis van taal en teksten

Ook wordt in de inleiding aangegeven dat als leerlingen op het (v)mbo onder niveau A2 scoren voor Nederlands dit voor een moedertaalspreker betekent dat hier mogelijk sprake is van ernstige leerproblemen en/of een cognitieve achterstand. Voor een tweedetaal spreker betekent het dat er een intensief NT2 traject nodig is om tenminste op niveau A2 te komen.

Bijlage 3 Raamwerk Rekenen/wiskunde

Het raamwerk rekenen/wiskunde is een instrument om:

- per beroepskwalificatie aan te geven welke vaardigheden op welke reken/wiskundige deelgebieden op welk niveau op een functionele manier beheerst moeten worden. Per mbo-kwalificatie kan in het dossier het niveau voor rekenen/wiskunde in de vorm van een profiel vastgelegd worden.
- het minimumniveau van rekenen/wiskunde vast te leggen dat nodig is voor het functioneren als burger in het kader van leren, loopbaan en burgerschap. Dit niveau wordt vastgelegd in het brondocument Leren, Loopbaan en Burgerschap. Dit document is een verplicht onderdeel van het mbo-onderwijs en zal dus geëxamineerd moeten worden.
- een gemeenschappelijk begrippenkader te creëren voor het beschrijven van reken/wiskundige vaardigheden voor alle beroepsopleidingen.
- individuele deelnemers in staat te stellen zelf in kaart te brengen wat ze op het gebied van rekenen/wiskunde beheersen, op een manier die voor alle betrokkenen herkenbaar is. Op die manier kan er gewerkt worden aan doorlopende ontwikkeling van de individuele deelnemers, vanaf het startniveau (intake, eerder verworven competenties) naar het beoogde uitstroomniveau van de gevolgde mbo-opleiding. De individuele deelnemer heeft altijd eigen sterke en zwakke punten en daarmee een eigen reken/wiskundeprofiel.

Het raamwerk rekenen/wiskunde is in eerste instantie gericht op het middelbaar beroepsonderwijs. Met deze doelgroep in gedachten is de vulling van de cellen geschreven. Omdat er sprake is van een beroepskolom van vmbo tot en met hbo zijn de niveaus ook voor deze onderwijssectoren van belang. Voor het vmbo gaat het dan met name om de onderste niveaus en voor het hbo om de bovenste niveaus. Het hoogste niveau in het raamwerk (Z2) overstijgt zelfs het mbo en is alleen relevant voor hbo (en wo). Dit niveau is opgenomen om de samenhang binnen de beroepskolom zichtbaar te maken.

Omdat het document onlangs (december 2007) in eerste versie is opgeleverd, ontbreken nog ervaringen met het document als het gaat om

- Het vastleggen van de beheersingsniveaus op de vier vaardigheidsgebieden voor de kwalificatiedossiers (acties Kenniscentra)
- Het geven van voorbeeldvullingen voor onderwijs en toetsing voor diverse beroepsprofielen
- Het vastleggen van de minimeisen voor de burgerschapscomponent
- Het ligt in de bedoeling om deze aspecten in 2008 nader uit te werken.

Voor het vaststellen van de eisen binnen de kwalificatiedossiers is nader onderzoek binnen de beroepspraktijk en ook binnen de beroepsopleiding zelf naar de aard en de verschijningsvormen van vereiste 'basisvaardigheden' rekenen/wiskunde nodig. Er wordt te makkelijk geroepen dat 'ze niet meer kunnen rekenen', maar ondersteuning van deze uitspraak door feitelijke gegevens over welke vaardigheden

dan niet (meer) worden beheerst ontbreken geheel. In het buitenland, met name Engeland, is uitgebreid onderzocht welke reken/wiskundige vaardigheden van essentieel belang zijn binnen de beroepspraktijk. Een algemene conclusie was dat de vaardigheden die binnen een beroepscontext zijn vereist weinig te maken hebben met de manier waarop in de school wiskunde werd geleerd. De transfer van geïsoleerde kennis en vaardigheden naar het functioneel gebruiken daarvan binnen een probleemsituatie binnen een beroepscontext had nooit plaatsgevonden. Een mooi citaat van een vooraanstaande Amerikaanse wiskundige (L. A. Steen, 2003) zegt in dit verband veel:

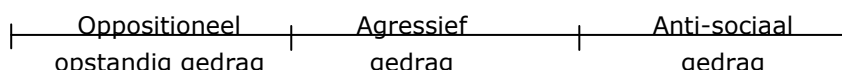
The contrast between these two perspectives—mathematics in school versus mathematics at work—is especially striking. Mathematics in the workplace makes sophisticated use of elementary mathematics rather than, as in the classroom, elementary use of sophisticated mathematics. Workrelated mathematics is rich in data, interspersed with conjecture, dependent on technology, and tied to useful applications. Work contexts often require multistep solutions to open-ended problems, a high degree of accuracy, and proper regard for required tolerances. None of these features is found in typical classroom exercises.

Bijlage 4 Inzet van schalingstechnieken voor examens

Een betrouwbaar, valide én flexibel model van examinering wordt bewerkstelligd door schalingstechnieken te gebruiken voor het meten van leervorderingen. De inzet van schalen voor het verrichten van metingen in het onderwijs is op zichzelf niet nieuw. Docenten die op systematische wijze het gedrag van hun leerlingen in kaart willen brengen, maken al heel lang gebruik van schalen. Vaak zijn dat eenvoudige schalen die de te meten objecten slechts benoemd in onderscheidende zin. Zoals bijvoorbeeld onderstaande (nominale) schaal om de leerstijl van een onderwijsdeelnemer te benoemen:



Iets ingewikkelder zijn de zogeheten ordinale schalen die de te meten objecten kunnen ordenen op grond van de te meten eigenschap, zoals bijvoorbeeld onderstaande schaal om de ernst van ongewenst gedrag te meten:

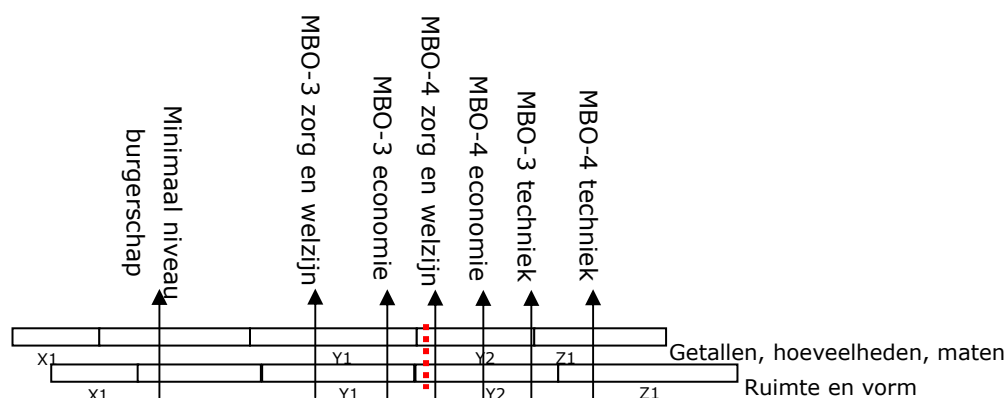


In beide schaaltypen kunnen de waarden van de te meten eigenschap ook met getallen worden aangeduid. Maar echt rekenen met deze getallen is niet goed mogelijk omdat er geen vaste afstanden tussen de opeenvolgende waarden zitten. Interessant wordt het pas vanaf de intervalschalen. Bij dergelijke schalen is er een meeteenheid die de schaal specificiert waarmee de te meten eigenschap numeriek gerepresenteerd wordt. Bij een lengteschaal, bijvoorbeeld, is die meeteenheid vaak de centimeter. Vanwege de betere mogelijkheden die dit biedt om prestaties van leerlingen onderling of met een of andere standaard te vergelijken, wordt er in het onderwijs ook gestreefd naar meten op intervalniveau. De moderne testtheorie biedt het instrumentarium om de leervorderingen van leerlingen uit te drukken in getallen op een intervalschaal.

De afgelopen jaren zijn voor onderwijs toepassingen al vaak intervalschalen geconstrueerd. Recente en aansprekende voorbeelden zijn de schalen die gebruikt worden ten behoeve van de verplichte instroomtoetsen rekenen en Nederlands voor pabo-studenten. Dergelijke schalen komen tot stand door te onderzoeken of de toetsopgaven, die een operationalisatie vormen van de leerdoelen van een bepaald vakgebied of deelvaardigheid daarbinnen, functioneren volgens een van tevoren geselecteerd model.¹ Toetsopgaven die zich 'gedragen' volgens het model worden definitief opgenomen in de opgavenbank. De opgaven in deze bank kunnen geordend worden naar moeilijkheidsgraad om zo een schaal te vormen voor het meten van de betreffende vaardigheid. Het schaalbegrip houdt in dat een kandidaat die een bepaalde opgave correct beantwoordt met een grotere kans ook correct zal antwoorden op opgaven met lagere schaalwaarden. Echter, naarmate opgaven met hogere schaalwaarden worden aangeboden zullen de kansen op een correct antwoord steeds verder afnemen.

De schaaltechnieken uit de moderne testtheorie bieden de volgende voordelen voor centrale examinering:

1. Empirisch gefundeerde informatie over de betrouwbaarheid. De kwaliteit van de examenopgaven is op voorhand onderzocht in een proefafname bij een representatieve steekproef uit de doelgroep. Alleen opgaven die voldoen aan de kwaliteitseisen worden opgenomen in de opgavenbank waaruit later het examen zal worden samengesteld. Met de beschikbare itemparameters kan voorafgaand aan het examen een schatting worden gemaakt van de meetnauwkeurigheid (en dus van de betrouwbaarheid) van het examen.
2. Empirisch gefundeerde informatie over de validiteit. De validiteit is, in ieder geval ten aanzien van één aspectⁱⁱ, gewaarborgd doordat in de reeds ter sprake gekomen proefafnames gekeken wordt of de opgaven functioneren volgens een bepaald model. De redenering is dat alle opgaven waarvoor het model geldt dezelfde achterliggende vaardigheid meten en dus dat het erg waarschijnlijk is dat de toetsen of examens die daarmee worden samengesteld constructvalide zijn; dat wil zeggen de eigenschap meten waarover men een uitspraak wil doen.
3. Transparante resultaten. De schalingstechnieken maken het niet alleen mogelijk om zowel opgaven (qua moeilijkheidsgraad) als kandidaten (qua vaardigheid) op één en dezelfde schaal af te beelden maar ook om de prestatiestandaard(en) daarop een plek te geven. Uitslagen in de vorm van schaalwaarden maken op een meer eenduidige wijze helder wat iemand kan en (nog) niet kan dan de vaak gebruikte schoolcijfers. Met name geldt dit voor schalen waarvan de onderliggende opgavenbanken een operationalisatie zijn van vakgebieden die zich uitstrekken over meerdere onderwijsniveaus (zie figuur 1).



Figuur 1: Voorbeeld van een MBO-vaardigheidsschaal voor rekenen/wiskunde.

Het onderbroken streepje geeft het examenresultaat aan voor rekenen/wiskunde van kandidaat X die een MBO-4 opleiding volgt in de sector Zorg en Welzijn. Het resultaat voldoet niet aan de gestelde norm, die voor beide deelgebieden ongeveer begin van niveau Y2 is gelegen. Er wordt wel voldaan aan de norm voor een MBO-3 opleiding in de betreffende sector en ruimschoots aan het niveau dat nodig is om als burger te functioneren in

de samenleving. Voor het deelgebied 'Ruimte en vorm' kunnen dezelfde conclusies getrokken worden.

4. Eenvoudige normhandhaving. De inhoud en de moeilijkheidsgraad van examens zijn van grote invloed op het leren van de onderwijsdeelnemers en het onderwijzen van de docenten en daarmee op het beheersingsniveau ten aanzien van het geëxamineerde vakgebied of deelvaardigheid daarbinnen. De bewaking van dat niveau vergt handhaving van de normen die bij examens worden aangelegd. In de examenpraktijk blijkt dat een heel lastige klus te zijn. Inhouden van examens kunnen vrij eenvoudig constant worden gehouden, maar voor de moeilijkheidsgraad ligt dat veel lastiger. Als de moeilijkheidsgraad van twee examenversies aanzienlijk van elkaar verschilt, is de consequentie meestal dat de normering wordt aangepast. Handhaving van de normering en dus van het onderwijsniveau is bij traditionele examens een probleem. Als examens worden samengesteld uit geschaalde opgavenbanken hoort dit probleem tot het verleden. De norm is dan niet langer gebonden aan een bepaalde examenversie maar ligt, na uitvoering van een normeringsprocedure, verankerd op de gebruikte vaardigheidsschaal.

ⁱ Er zijn verschillende modellen waarvan er hier één wordt beschreven. In dat model is de moeilijkheidsgraad van een item gedefinieerd als de vaardigheid waarbij de kans op een correct antwoord precies 50% is. Een voorbeeld kan dit verduidelijken. Als we bij hoogspringen de lat exact zo hoog leggen als de sportman of -vrouw kan springen (dat wil zeggen op een hoogte die het gemiddelde is van alle lathoogtes waar die persoon overheen gesprongen is), mogen we verwachten dat hij of zij in de helft van alle gevallen erover zal springen en in alle andere gevallen de lat eraf zal springen. Als de moeilijkheidsgraad van het item groter is dan het vaardigheidsniveau van de persoon wordt de kans op een correct antwoord kleiner dan 50% (de lat wordt er vaker afgesprongen dan dat de sporter erover heen springt). Is de moeilijkheidsgraad van het item kleiner dan het vaardigheidsniveau dan wordt de kans op een correct antwoord groter dan 50% (de sporter springt vaker over de lat dan dat hij deze eraf springt). Dit proces wordt in een wiskundig model beschreven waarin de kans gespecificeerd wordt op het geven van een goed antwoord door een persoon met een bepaalde vaardigheid. Deze kans is uiteraard afhankelijk van itemkenmerken zoals de moeilijkheidsgraad en het discriminerend vermogen.

ⁱⁱ Er worden in de relevante literatuur verschillende bewijsbronnen voor validiteit onderscheiden: inhoudsgerelateerde, constructgerelateerde en criteriumgerelateerde bewijzen voor validiteit.